In-silico screening of potential inhibitors of gamma-secretase, a key enzyme of Alzheimer's disease

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Technology

in

Biotechnology

by

ARUN E V R

Department of Biotechnology and Medical Engineering

National Institute of Technology, Rourkela

Rourkela, Odisha 769008, India

May 2013

In-silico screening of potential inhibitors of gamma-secretase, a key enzyme of Alzheimer's disease

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Technology

in

Biotechnology

by

ARUN E V R

Department of Biotechnology and Medical Engineering

National Institute of Technology, Rourkela

Rourkela, Odisha 769008, India

May 2013

Acknowledgement

I would like to express my gratitude to my guide, Dr. Nandini Sarkar, without her guidance the project would not have shaped up to what it is. I would not be exaggerating the freedom she had given me with regards to the design of the project. Working under her has been one of the best learning experiences I have had. I would also like to thank the Department of Biotechnology and Medical Engineering for giving me the space to carry out this project. Also I would like to acknowledge the support of the faculty of the Department of Biotechnology and Medical Engineering, especially Dr. A. Thirugnanam.

I extend my gratitude to Dr. Bibekanand Mallick, Department of Life Sciences, NIT – Rourkela, for his valuable technical guidance. My gratitude is also towards Dr. Prabodh Borah and his team, Assam Agricultural University, Khanapara, Guwahati, for accepting my request for participation in the training titled "Basics of Molecular Modeling and Computer-Aided Drug Designing". The participation in the workshop has been crucial for my learning of techniques implemented in this project.

I would like to thank Mr. Suresh Kumar A, Department of Chemical Engineering, NIT, Rourkela, for his suggestions.

Also I would like to thank my family, frineds and colleagues, especially, Mr. Ritiritao Brahma, Mr. Om Shankar Awasthy and Mr. Akshaya Kumar Padhi for their support.

Abstract

Gamma-secretase is a trans-membrane aspartyl protease that consists of four subunits, namely Anterior Pharynx Defective Phenotype (APH-1), Presenilin (PSEN), Nicastrin (Nct) and Presenilin 2 Enhancer (PEN2). Presenilin is identified as the catalytic core of gamma-secretase with the two aspartyl residues at the catalytic site. Gamma-secretase is involved in the ultimate step in the processing of Amyloid Precursor Protein (APP) to yield Amyloid Beta peptide (Aβ). Aβ of various residue lengths is formed that includes the toxic Aβ42. Aggregation of Aβ42 has been identified to contribute to etiology of Alzheimer's disease (AD). Inhibition of processing of APP by gamma-secretase is a possible intervention strategy for the therapy for AD. Various studies carried out to find inhibitors for gamma-secretase have populated two classes of compounds, Gamma-secretase Inhibitors (GSI) and Gamma-secretase Modulators (GSM). Despite these efforts there is a dearth of an effective therapy for AD. This study aimed to find potential inhibitors of gamma-secretase using in-silico screening of DrugBank database. A total of 10 Pharmacophore models were developed from 54 molecules shown to inhibit gamma-secretase. The pharmacophore models were used as 3D query in database screening of 6160 drug molecules selected from DrugBank. The list of hits that resulted from the database screening using all the 10 pharmacophore models was compacted to yield 721 unique entries with a fit score over 3.00 on a scale of 4.00 against the pharmacophore model. A QSAR model was developed employing multiple linear regression to calculate the predicted IC50 value for the 721 molecules from screening using pharmacophore models. Docking study was done to calculate the binding energy for 498 molecules with predicted IC50 value under 10000nM. 55 molecules with binding energy, ΔG, lesser than –8.00 kcal/mol are presented as potential inhibitors of gamma-secretase. Thus, this data can be used for further studies for the development of therapy for Alzheimer's disease.

Table of Contents

List of Figures

List of Tables

1 Introduction

Proteins are the cogs that make the cell machinery tick. Proteins are involved in every aspect of cell function namely structure, motion, catalysis, recognition and regulation. The cellular membrane is embedded with membrane proteins that act as receptors, channels for a variety of molecules. These proteins thus enable signal transduction and transport of molecules into and out of the cell. Enzymes are proteins that act as catalysts in many processes and pathways in the cell. Also a number of enzymes are known to regulate various cellular processes and pathways by activating and deactivating other molecules involved in the processes. Antibodies are also proteins that are part of the immune response system that help to maintain immunity against foreign bodies. Signal transduction occurs due to the participation of trans-membrane proteins that act as receptor channels for the transport of the signaling molecules across the cell membrane. These are a few of the innumerous class of proteins involved in the complex machinery in the cell.

1.1 Protein Folding and its regulation

Self-assembly is one of the most intricate processes of the living system. A classic example of this self-assembly is the folding of proteins into three-dimensional structures. The bringing-together of the functional groups in vicinity to one-another in a highly specific structures result from protein folding aid in the possibility of diversity as well as selectivity in the chemical processes. Many biological activities are coupled to the correct folding of proteins that include molecular trafficking, regulation of growth and differentiation of the cell. Thus only a correctly folded protein has a long-term in the biological environment. The most thermodynamically structures of proteins always correspond to the native state of the protein(1). It is established that folding of a protein

does not involve the systemic search of all the possible conformations but a stochastic search of the conformations accessible to the chain of polypeptide(2). The right energy landscape, encoded by the amino acid sequence of the protein, which is in turn, is determined by the genetic code that has evolved for rapid and efficient folding of the protein, enables the scan of a small number of all possible conformations in order for the coil to transition into a native structure. This energy landscape approach to protein folding shows that the transition states for folding are critical regions of energy surfaces through which all molecules must pass to reach the folded native state(3).

The results of many studies show the formation of nucleus from a small number of residues, over which the rest of the structure condenses, as the fundamental mechanism of protein folding(4). An important element in the folding process is the formation of secondary structure that is stabilized by hydrogen bonding between the amide and carbonyl groups of the peptide main chain. Studies involving the structural properties of the intermediate or transition states show that a full native structure is obtained only when native-like interactions both within and between the domains have been formed(2). This is achieved by close packing and locking of side chains and exclusion of water molecules from protein core.

Folding of proteins in vivo has been established to be co-transitional in a few cases, which implies that the protein synthesis and folding occur simultaneously. In other cases the proteins fold in the cytoplasm or after trafficking and translocation through the membranes in mitochondria or the endoplasmic reticulum. In order to regulate the folding of proteins several molecular chaperones present in the cell are brought into play. These molecular chaperones interact with the nascent chain of protein as soon as they emerge from the ribosomes or later. Molecular chaperones do not act as catalysts for the steps in protein folding but increase the effectiveness of these steps by reducing the probability of

the competing reactions such as aggregation. The mechanism of action of the chaperones can be illustrated with the chaperonin GroEL and its co-chaperone Gro-ES. These have been shown to protect the protein from the external environment by allowing the incompletely folded protein to enter its cavity and then fold into native structure(5). It is also suggested that molecular chaperones are able to rescue misfolded proteins and aggregated proteins enabling them to have another chance at folding correctly. Also available are a class of folding catalysts such as protein disulfide isomerases, peptidylprolyl isomerases(6). The proteins destined for secretion undergo a quality control check in the endoplasmic reticulum. The mechanism of quality control involves glycosylation and deglycosylation reactions to enable correctly folded proteins to be distinguished from misfolded proteins. The incorrectly folded proteins initiate an 'unfolded protein response' in which they are ubiquitinated and then degraded by proteasomes in the cytoplasm.

1.2 Protein Misfolding and Amyloidogenesis

It is understood that proteins need to be folded correctly to survive in the cell and carry out its function. It is also evident that both folding and unfolding are taken as ways of generating and abolishing cellular activities. The precursor to the protein degradation is unfolding(7). An intrinsic property of proteins is misfolding that happen continuously and is influenced by the sequence of amino acids. .Mutations in the amino acid sequence are said to accelerate misfolding. In addition, environmental changes such as increased temperature, low or high pH, presence of oxidative reagents or high glucose level cause the loss of native conformation in proteins rapidly, while being termed denaturation. This results in unfolding of the protein. The unfolded proteins are not functional. Also the unfolded state is thermodynamically unstable and unfavorable. In an attempt to seek lower energy level and higher stability, these proteins have aggregation propensity. Aggregation

starts with nucleation in which the proteins attach reversibly onto a core that is growing, termed the nucleus, followed by irreversible attachment of proteins once the critical mass of the nucleus crosses the threshold to form a larger aggregate. The concentration of the monomers also determines the onset of aggregation. When the concentration of the monomer is low, the monomeric state is favored while at high concentration the aggregated state is favored as it presents a large energy barrier for resolution (8). Thus the aggregates are highly stable. Protein segments that have hydrophobic residues, low net charge and susceptibility to β-sheet formation are thought to initiate aggregation (9), by acting as the precursor cores that facilitate aggregation further. The precursor is generated from the native structure by means of introduction of misfolded variant, incorrect or incomplete proteolysis. Thus, there is introduction of a precursor pool from the partial misfolding of the variant protein from which rapid aggregation occurs.

The structure of the protein aggregates vary from unordered amorphous aggregates to highly ordered fibrils. These highly ordered fibrils are termed as amyloid. The amyloid conformation is stable yet reversible. The amyloids are rich in cross-β structure. It consists of non-branched, linear fibrils of protein or peptide. X-ray analysis revealed that the β strands are arranged perpendicular to the fibril axis(10). Recent studies reveal the presence of parallel(11) or antiparallel(12) β-sheet conformation. Thus, confirming that amyloids fibrils have no universal tertiary and quaternary structures. Amyloids show conformational plasticity by adopting more than one stable tertiary fold due to which conformational differences exist within the fibril formed from a single polypeptide.

Amyloid proteins are believed to be non-functional proteins while in some instances have been associated with physiological function, for example, fibrils of curlin in *Escherichia coli* mediate binding of bacterium to host. A large number of proteins, tau,

amyloid β, prion, insulin, α-synuclein and lysozyme to name a few, have been identified to undergo misfolding to form amyloids and result in misfolding diseases.

1.3 Amyloidoses

The group of protein misfolding diseases where there is accumulation of protein aggregates systemically or locally in certain tissues or organs is termed amyloidosis. The characteristic feature of the amyloidoses is the presence of plaques or inclusion bodies. The conditions for the development of the disease are not yet fully studied. These diseases can be systemic or localized based on the location of occurrence of the protein aggregates. The symptoms depend on the protein involved and the site or organ or aggregation of the protein. Diagnosis of the diseases requires histological analysis followed by the definition of the amyloidosis. Efficient therapy is not available for most of the amyloidoses.

It is evident that the aggregates from proteins and peptides can be toxic to the cells or deviate from the function of the native protein. The amyloid fibrillar aggregates were earlier thought to be involved in the pathogenesis of a number of amyloidoses. Evidence has been piling-up on the involvement of the prefibrillar aggregates(13), oligomers(14) and not-native monomers(15) in exhibition of toxicity in the cells or tissues. Despite the lack of complete mechanism of various diseases and the cause for toxicity to cells or tissue, it can be said that the conversion proteins from soluble to insoluble state results in a population of non-native structures. The toxicity of these misfolded species tends to arise from the fact that the hidden groups that are normally buried are exposed and dispersed. Small aggregates are relatively more toxic that mature fibrils as these have a higher proportion of surface residues(16). These non-native structures are supposed to trigger a number of uncharacteristic events on its interaction with the various molecules and components of the cell. Hence are likely to cause malfunction in the cell machinery resulting in the diseased condition. Also the damage to the architecture of the tissue brought forward by

the accumulation of the aggregates is supposed to be the reason of pathogenesis in a few diseases(17).

A number of human diseases have been classified as amyloidoses namely Alzheimer's disease, Parkinson's disease, Huntington's disease, Dementia with Lewy bodies, lysozyme amyloidosis, fibrinogen amyloidosis, AA amyloidosis, AL amyloidosis, cataract and cutaneous lichen amyloidosis.

1.4 Alzheimer's disease and gamma secretase

Alzheimer's disease is an irreversible, progressive brain disease that slowly destroys memory and thinking skills, and eventually even the ability to carry out the simplest tasks. In most people with Alzheimer's, symptoms first appear after age 60. It was estimated that 35.6 million around the world may have Alzheimer's disease in 2010; 65.7 million by 2030 and 115.4 million by 2050(18). Alzheimer's disease is the most common cause of dementia among older people. Dementia is the loss of cognitive functions which ranges in severity from mild to aggressive. In 1906, Dr. Alzheimer noticed changes in the brain tissue of a woman who had died of an unusual mental illness. Her symptoms included memory loss, language problems, and unpredictable behavior. After she died, he examined her brain and found many abnormal clumps (now called amyloid plaques), tangled bundles of fibers (now called neurofibrillary tangles) and the loss of connections between nerve cells (neurons) in the brain. Hence, Alzheimer's disease is named after Dr. Alois Alzheimer.

In Alzheimer's disease, senile plaques are formed by accumulation of aggregates of 40- to 42-amino acid long, 39- and 43-aa peptides have also been described, amyloid-β (Aβ) peptide. Amyloid-β is formed by cleaving amyloid precursor protein (APP) by β- and γ-secretases. The cleavage of APP by γ-secretase produces a 42 amino acid Aβ-42 or a 40 amino acid Aβ-40 is formed when it occurs in endoplasmic reticulum or trans-Golgi

network respectively. APP is involved in various physiological functions while that of Aβ peptide have not yet been deduced. In 1991, David Allsop and John Hardy proposed the amyloid cascade theory (ACH), according to which the amyloid-β peptide aggregate into protein assemblies, which in turn deposit into plaques in the brain that cause neurotoxicity(19).

1.5 Therapy for Alzheimer's disease

The current treatments for Alzheimer's disease are mostly symptomatic, while disease-modifying agents are emerging. The symptomatic or palliative treatments attempt to compensate for the decreased activity of cholinergic neurons. Various classes of compounds such as choline donors (4-Aminopyridine), cholinergic agonists (nicotine analogs) and acetyl choline esterase (AChE) inhibitors (donezepil, rivastigmine, galantamine) have been developed. The disease modifying drugs are developed to decrease the concentration of Aβ. This is done by inhibition of β- and γ- secretases. Also Aβ can be degraded or excreted and cleared from the cell to decrease its concentration in theory. In another approach the aggregation of Aβ can be inhibited. Inhibition of tau protein's hyperphosphorylation by inhibiting the enzymes involved namely tau protein kinase I and II, extracellular signal-regulated kinase. Also drugs are also being developed that inhibit excitotoxicity (propiracetam, oxiracetam, or aniracetam) or act as antioxidants (trans-resveratol, EGb761, Clioquinol). Thus there is a need for the increased effort in the development of therapy for Alzheimer's disease.

1.6 In-silico screening in drug discovery

Drug discovery process consists of several steps namely target identification, lead compound screening, lead optimization, ADMET studies, preclinical trial evaluation, clinical trials and registration. This process trails behind due to being time-consuming,

expensive and inefficient due to the low rate of novel therapeutic discovery. In order to increase the overall efficiency of the process new methodologies and technologies are being implemented. Computational methods are employed in various areas of research. It is estimated that the implementation of computational methods in drug discovery can cut the annual cost by 33% and the time required by 30% for developing a new drug. Computational techniques are capable of providing information regarding function prediction, pathway information, homologue mapping, structural information, chemical information and disease association. Also a vast number of resources are available such as databases of chemical molecules such as ChEMBL, Therapeutic Target Database (TTD), Potential Drug Target Database (PDTD), PubChem, ChEMBLdb, DrugBank and etc.

In-silico screening or virtual screening has become an established technique for the development of hit lead compounds. Virtual screening complements traditional high-throughput screening (HTS) assays. This technique can act as a filter by eliminating a vast majority of non-binding compounds in-silico thereby reducing the number of compounds to be tested in vitro, thus, reducing the investment of resources in the early stages of drug discovery. It includes a series of techniques such as simple filtering and pharmacophore searches to docking and scoring. In-silico screening starts from a database of real compounds or a virtual database.

1.7 Pharmacophore modeling

A pharmacophore can be described as set of steric and electronic properties required for a compound to bind to an active site. It describes the three-dimensional arrangement of molecular features. The components of a pharmacophore model are hydrogen bond donors, hydrogen bond acceptors, aromatic rings, hydrophobic centers, acidic groups, basic groups, planar atoms, positive charge centers, negative charge centers, etc.

Pharmacophore modeling is the identification of various components of the pharmacophore and definition of the components to forma model. Pharmacophore modeling can be divided into structure-based and ligand-based approaches. Structure based pharmacophore modeling is construction of model from available structural data. It can be further divided into macromolecule-ligand complex based and macromolecule based. Thus a pharmacophore model can be developed from a set of active compounds without any knowledge of the protein active site or from the active site of the protein when the geometry of the active site is available. This model can be used to search a compound database, which can hence be screened.

1.8 Quantitative Structure-Activity Relationship (QSAR)

QSAR is a way of finding a simple equation that can predict an unknown property from the molecular structure of a compound. This is done by curve fitting to find the equation coefficients that are weights for known molecular properties termed descriptors. A descriptor can be any number that describes some aspect of the molecule. The property that is being predicted is termed as activity. A QSAR model is essentially an equation that can be developed by various methods such as multiple linear regression, partial least squares, linear discriminant analysis, logistic regression, genetic function approximation, k-nearest neighbor, feed forward back propagation neural network, general regression neural network and support vector regression.

1.9 Docking method

Docking is the search for the energetically favorable binding pose of a ligand to a macromolecular receptor. The aim of the docking program is to accurately predict the structure of the ligand-receptor complex enabling the prediction of the binding free energy of the complex. Docking has enabled the search through databases of available chemical

compounds for ligands to use as therapeutics or in studies. The molecule preparation must be performed before docking; it involves the creation of complete molecular model from molecular structure file followed by processing of the structure for use in specific docking program. The sampling method generates multiple positions of the ligand within the receptor such that the ligand is placed using the constraints of the receptor shape and are near minimum-energy positions. The three classes of sampling method are rigid ligand docking, flexible ligand docking and flexible receptor docking. The scoring method is used to compute the binding energy of a ligand to the receptor. The accuracy of the scoring function inversely affects the speed of calculation. Thus it is necessary to find a balance in choosing the method to be employed.

1.10 Objective of the study

The major objectives of the study are

- To find the potential inhibitors of gamma secretase by screening of a chemical database
- To develop pharmacophore models for use a query to screen the chemical database
- To compute the specific activity of the hits of database screening using pharmacophore models
- To subject a set of compounds filtered based on the predicted specific activity to docking study against the receptor, here gamma secretase
- To put forward a set of compounds filtered based on binding energies computed using docking study as potential inhibitors of gamma secretase

2 Literature Review

Alzheimer's disease is a neurodegenerative disorder where the affected individuals exhibit dementia. Affected individuals exhibit progressive and permanent decline in memory and cognitive functions. It is difficult to positively diagnose Alzheimer's disease in the early stages as memory loss can have other causes too. However a consistent pathology with amyloid aggregates called 'plaques' and neurofibrillary tangles, is observed in the brain of individuals affected by Alzheimer's disease(20). These are found in the temporal neocortex and hippocampal regions of the affected individual's brain(21). Of the different forms of Alzheimer's disease the most common form occurs after 6o years of age and is sporadic(22). The other form of AD, Familial Alzheimer's disease (FAD) has onset before 60 years of age (22). The brain regions involved in learning and memory process of the individuals with AD are reduced in size. The areas for episodic memory, attention, executive functions, semantic memory, language and spatial orientation deteriorate as the disease progresses(23, 24).

The amyloid plaques and neurofibrillary tangles result from the deposition of amyloid-β peptide and hyperphosphorylated tau protein, which lead to neuronal loss and neurotoxicity in the affected brain. The mechanism of induction of neurotoxicity by the deposition of Aβ peptide is not clear. Oxidative stress is stimulated by Aβ peptide directly and indirectly. Aβ peptide reduces metal ions to produce hydrogen peroxide by acting as enzymes and also generating free radicals (25-27). Aβ peptide also produces free radicals by binding to mitochondrial proteins (28). Aβ peptides generate oxidative stress through neuroinflammation. Astrocytes and microglia are activated and participate in the immune response to the deposition of amyloid beta peptide. These brain cell types produce reactive oxygen species upon activation in addition to chemokines and cytokines leading to cell

death(29). Aβ peptide also stimulates microglial cells to produce reactive nitric oxide, by inducing the expression of nitric oxide synthase(30), and quinolinic acid, a neurotoxin(31).

Aβ peptide is produced by the processing of amyloid precursor protein (APP) by proteolysis. It then accumulates and eventually deposits into amyloid plaque in the brain. This is a primary event in the pathogenesis of AD (22). Along with the secondary events namely activation of microglia and astrocytes and formation of neurofibrillary tangles the formation of amyloid plaques result in neuronal dysfunction and loss. This hypothesis is termed as amyloid cascade hypothesis (ACH)(32). APP is a type I integral membrane protein that is expressed widely. It is expressed by the APP gene located on chromosome 21. It is produced in different isoforms from 695 to 770 amino acids. It is processed by secretase enzymes to produce Aβ peptide in two distinct pathways, amyloidogenic and non-amyloidogenic. In amyloidogenic pathway the cleavage of soluble APP by β-secretase gives soluble β-APP, which are released in to the extracellular matrix(33). It is followed by the cleavage of the C-terminal fragment (C99) by γ-secretase which releases Aβ peptide of varying lengths with Aβ-40 and Aβ-42 being the most common forms. In non-amyloidogenic pathway, soluble APP is cleaved by α-secretase to release α-APP into the extra-cellular matrix. The cleavage of the resulting C-terminal fragment (C83) by γ-secretase produces p3 fragment(34). Aβ-42 exists as a monomer, which undergoes oligomerisation to form oligomers, proto-fibrils and fibrils. Aβ-42 aggregates into β-sheet fibrillar state or non-β-sheet non-fibrillar state(35, 36). The oligomeric intermediate species and the mature fibrils are neurotoxic(37).

Research has identified the genes responsible for the pathologies namely the amyloid precursor protein (APP) and the presenilins, PS1 and PS2 (PSEN1, PSEN2)(38). PS1 is central to the γ-secretase complex that cleaves APP and Notch(39). Mutations in these genes have been identified in Familial Alzheimer's disease (FAD). These account for

50% of the FAD cases(40).The pathological features of FAD is similar to sporadic AD, hence mutations in APP, PSEN1, PSEN2 may have a role in sporadic AD too. So far 33 mutations in APP, 185 mutations in PSEN1, 13 mutations in PSEN2 have been documented at the Alzheimer Disease & Frontotemporal Dementia Mutation Database(41) (http://www.molgen.vib-ua.be/ADMutations).

γ-secretase is a transmembrane aspartyl protease. It exists in the plasma membrane as a mature complex. It is a multiprotein complex consisting of Presenilin, Nicastrin, Aph-1 and Pen-2. It may possess homogeneous activity, as several alternatively spliced variants and paralogs of Presenilin and Aph-1 have been identified, or heterogeneous activity. The presenilins (PS) form the active catalytic core of the enzyme. PS1 and PS2 are the homologs of presenilin that exist in mammals. These have a molecular weight around 50kDa. It spans the cellular membrane several times. It has been identified that two aspartyl residues Asp 257 and Asp 385 located on transmembrane domains 6 and 7 respectively are central to the catalytic activity of the enzyme(42). Presenilins are synthesized as precursor proteins that are incorporated into a larger complex for stabilization. This is accompanied by a proteolytic cleavage of PS by presenilinase(43). This results in the formation of two separate fragments N-terminal fragment (NTF) and C-terminal fragment (CTF). Each has a molecular weight of ~ 30kDa and ~20kDa respectively. The second member of the complex Nicastrin (Nct) is a glycosylated membrane protein that binds to PS. It is synthesized as a precursor that is transported by PS from the endoplasmic reticulum and is required for the stabilization of PS NTF and CTF(44). Aph-1 is a multimembrane spanning protein with a molecular weight ~30kDa. It is necessary for the transport of Nct to cell surface(45). Pen-2 is a hairpin-like membrane protein with a molecular weight of approximately 12kDa(46). The molecular weight of the γ-secretase complex is on debate, with estimates from 200-250kDa(47) to ~440kDa(48)

and larger. All the four proteins are necessary for the full proteolytic activity of the enzyme. Genetic studies or knockdown studies on one or more of the components caused the loss of activity of the enzyme. Overexpression of one or more of the proteins but not all had no effect on the activity of the enzyme. Overexpression of all the component proteins caused a change in the proteolytic activity (48-50).

In an attempt to develop therapy for Alzheimer's disease, the proteolytic cleavage of APP by γ-secretase has been targeted in various studies. This led to the development and discovery of a number of compounds that inhibit or modulate the activity of γ-secretase. Gamma-secretase inhibitors inhibit the activity of gamma-secretase on not only APP but other substrates as well. Administration of potent gamma-secretase inhibitors to APP transgenic mice resulted in gastrointestinal toxicity due to the disruption of Notch signaling in the ileum(51). Thus the development of compounds that are specific for the cleavage of APP was necessary. This gave rise to the class of compounds called gamma-secretase modulators which spared the cleaving of Notch by gamma-secretase.

γ-secretase inhibitors include compounds of various classes. These include transition state analogs, dipeptidic inhibitors, sulfonamides, kinase inhibitor, NSAIDs(52). L-685,458(53), WPE-31-C, LY450139(54), DAPT(55), Imatinib(56), R-flurbiprofen(56), sulindac sulfide(57), AZ1136, AZ3303, AZ4800(58), E2012(59), MRK-560 are the compounds that belong to the mentioned classes. Also studies have shown potent γ-secretase inhibitors that include α-hydroxycarbonyl derivatives(60), aryl sulfones(61), compounds based on benzobicyclo[4.2.1]nonane core(62), helical D-tridecapeptides(63), derivatives of triterpene glycoside(64, 65), fused oxadiazines(66), derivatives of piperidine acetic acid(67), azepine derivatives(68) and cyclohexanones(69). The listed are only a segment of the compounds that have been reported. The mechanism of inhibition of

many of these compounds has not yet been fully understood. Meanwhile the search for new drugs utilizing various techniques is continuing.

Drug discovery and development consists of several steps. It starts with the pre-discovery, which is necessary to understand the pathogenesis of the disease and the pathways involved. It is a knowledge acquiring process which is crucial for treating the problem. Then target identification follows in which the molecule that is involved in the disease is identified as target that can be affected by a drug molecule. Finally, drug discovery begins with the testing of various molecules to find a lead molecule. This has been carried out by finding compounds in nature, de novo synthesis of molecules, high-throughput screening. Lead compounds are tested for Absorption, Distribution, Metabolism, Excretion and Toxicological properties. Lead compounds are optimized to better its properties. This is done by testing analogues of the leads, leading to the choosing of the candidate drug. Also considered are the formulation, delivery mechanism and large-scale manufacturing. Pre-clinical testing of the candidate drug involves in-vitro and in-vivo tests these include testing on animals. The clinical trials take place in four phases. In the first phase the drug is tested for safety in humans by testing on healthy human volunteers. The drug is tested on a small group of patients in the second phase of clinical trials. The third clinical trial phase is extensive where the safety and efficacy of the drug are tested. Then the drug is registered for approval with the authority for drug approval. It is followed by the final clinical trial phase where it is prescribed to a larger section of the population.

Computational methods have been taken up for the drug discovery process. This has resulted in the faster and less expensive drug discovery. The availability of cheminformatics resources such as chemical compound databases, target structure databases, in addition to the computational techniques such as molecular modeling, molecular simulation, etc. has favored the trend of computational drug discovery. One of

the most time and resource consuming processes in drug discovery is screening for hits and lead compounds. The advent of virtual or in silico screening aimed at changing that. An integral part of drug discovery of late is virtual screening. Virtual screening was introduced recently compared to the use of database searching. It was defined as 'automatically evaluating very large libraries of compounds' using computer programs(70). This suggests that a library of $>10^{60}$ compounds can be filtered to manageable number using in-silico or virtual screening, while the focus in practical scenarios is on filtering combinatorial libraries of targeted compounds and enriched libraries. The success of an in-silico screen lies in finding new or interesting structures than many hits. Availability of datasets, in-silico screening techniques and successful screening techniques has helped this method gain attention. The possibility of rapid identification of novel ligands(71), biomolecules from natural products(72) have improved this interest.

Virtual screening can be structure-based(73) or ligand-based(74). Ligand based approach is primarily used when the protein structure is absent, while structure-based approach is used in the presence of high atomic resolution structural data of the protein. Results of various studies indicate that structure-based and ligand-based approaches can yield a better result when applied in concert(75, 76). Recently, the successful screening for the target MEK1 was reported. It employed both in silico and in vitro methods to identify fragments. This result has been suggested for use in drug discovery project(77). In another study, a library of small molecules was screened to find drugs that can be combined with highly active antiretroviral therapy (HAART) to reactivate dormant viruses. Similarity based search was employed to the hits of the screen resulting in the find of 8-methoxy-6-methylquinolin-4-ol (MMQO) as a successful hit(78). Docking, fragment based search are structure-based approaches, while pharmacophore search, QSAR are ligand based approaches of in silico screening.

Pharmacophore was introduced and defined by Paul Ehrlich in 1909 (79) as ' a molecular framework that carries the essential features responsible for a drug's biological activity'. It is defined by IUPAC(80) as 'an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with the biological target and to trigger (or block) its biological response'. Superposing a set of active molecules and extracting the common chemical features required for bioactivity or probing the interaction points between the ligand and the macromolecular target can give a pharmacophore model. It is being extensively used in in silico screening as well as de novo design.

Structure-based and ligand-based are the two approaches for pharmacophore modeling. Ligand based pharmacophore modeling facilitates the drug discovery in the absence of the macromolecule structure. Pharmacophore generation from multiple ligands, termed training set, involves creating the conformational space for each ligand in the training set to represent conformational flexibility of ligands, and aligning the multiple ligands in the training set and determining the essential common chemical features to construct pharmacophore models. Structure based pharmacophore modeling involves the use of 3D structure of the macromolecular target. Its protocol starts with the analysis of an analysis of the complementary chemical features of the active site and their spatial relationships, and a subsequent pharmacophore model assembly with selected features. It can be further classified into macromolecule–ligand-complex based and macromolecule (without ligand)-based. A number of pharmacophore generator programs are available namely GASP(81), DISCO(82), HipHop(83), Hypogen(84), GALAHAD, PHASE(85), MOE, LigandScout(86), Pocket(87) and GBPM(88) besides a number of academic programs. On modeling of a pharmacophore model from a set of active ligands it can be used a template to query a 3D chemical database and search for potential ligands termed

hits. This approach faces the challenge of higher false positives and or higher false negative. This can be overcome by model optimization by adjusting the tolerance radius and position of each pharmacophoric feature.

QSAR stands for Quantitative Structure-Activity Relationship that describes structure-activity relationship in terms of steric properties and physiochemical parameters or certain structural features. Thus QSAR can correlate inhibition constants, rate constants, and affinities of ligands to binding sites with in a series of compounds. The properties of the compounds are termed as descriptors. Lipophilicity, polarizability are molecular descriptors. Enzyme inhibition data of the compounds has been studied by correlating with physiochemical properties using QSAR. QSAR is based on the assumption that different structural properties contribute in a linear additive manner to its biological activity. There are several methods available to develop a QSAR model the most common being multiple linear regression. In an investigation of inhibition of dihydrofolate reductase (DHFR) from *E. coli* and *L. casei* by benzylpyrimidines, two equations (QSAR models) were generated one for each organism. These models differed on the contribution of five substituents of benzyl group to biological activity. This was explained to be due to the presence of rigid leucine residue on *L. casei* DHFR hence it formed a narrow cleft that the methionine residue in *E. coli DHFR(89).*

Molecular descriptors have been extensively used in deriving structure-activity relationships, quantitative structure activity relationships, and machine learning prediction models for pharmaceutical agents. A descriptor is "the final result of a logical and mathematical procedure which transforms chemical information encoded within a symbolic representation of a compound into an useful number or the result of some standardized experiment". Many programs e.g. PaDEL-descriptor, DRAGON, Molconn-Z, MODEL, Chemistry Development Kit (CDK), JOELib and Xue descriptor set are

available to calculate physical and chemical descriptors. These methods can be applied to derive >3,000 molecular descriptors. These descriptors include constitutional descriptors, topological descriptors, RDF descriptors, molecular walk counts, 3D-MoRSE descriptors, BCUT descriptors, WHIM descriptors, Galvez topological charge indices and charge descriptors, GETAWAY descriptors, 2D autocorrelations, functional groups, atom-centered descriptors, aromaticity indices, Randic molecular profiles, electrotopological state descriptors, linear solvation energy relationship descriptors, and other empirical and molecular properties.

Docking and scoring are used to identify potential lead candidates. The algorithms generate a subset of a collection of compounds with higher affinity against a target macromolecule by predicting their binding mode and affinity. Docking algorithms deal with the prediction of ligand conformation and orientation with in the specific binding site of the receptor. DOCK(90), FlexX(91), GOLD(92), Glide(93), ICM(94), FRED(95) and AutoDock(96) are a few of the docking programs available in addition the web-based docking servers. Scoring methods attempt to estimate the bound receptor-ligand complex rightness. Scoring methods make assumptions and oversimplify various physical characteristics. This results in the inability of the scoring function to differentiate between correct and incorrect poses. Thus giving rise to a number of false positives and false negative hits. Docking based screening has been successfully implemented in finding the inhibitor for Acetylcholinesterase(97). The initial screening was done for the ACD and Maybridge database consisting of 160,000 compounds. The computer program used was ADAM & EVE.

3 Methods

3.1 Dataset Collection

The choice of training set is important to the pharmacophore generation process as the pharmacophore built is as good as the input data. The following criteria have been considered during the selection of data set. All the compounds have to bind to the same receptor more or less in the same fashion. The compound data should be widely populated over a larger activity range. The most active compounds have to be included in the dataset. All biologically relevant data should be obtained for the compounds(84). Each individual feature in the resulting pharmacophore will contribute a certain weight proportional to its contribution to biological activity. Taking in to account these criteria, 369 gamma-secretase inhibitors were sourced from various literature sources and the target database at ChEMBL (60-63, 65, 66, 98-109). The 2D structure of the compounds were obtained from ChEMBL chemical database(103). 2D structures were converted to 3D structures using OpenBabel 2.3.1(110). Next, the compounds were filtered based on No. of violations of Lipinski's rule of five=0 which yielded 206 compounds. Also on filtering compounds with IC50 value < 100nM in a range of 0.08nM to 93nM, 56 compounds were obtained. These 56 compounds were chosen as training set. The chemical structure of the compounds in the training set is given in (Figure 1), while their experimental activity (IC50) is given in (Table 1)

3.2 Feature Mapping

Pharmacophore features are chemical features such as hydrogen bond donors, hydrogen bond acceptors, aromatic rings, hydrophobic centers, acidic groups, basic groups, planar atoms, CO_2 centroid, NCN^+ centroid, etc. Pharmacophore features were mapped for 56 compounds of the training set using the Feature Mapping protocol in

Discovery Studio 2.5(111) The total no. of pharmacophore features among the compounds ranged from 8 to 28.(Table 2)

3.3 Pharmacophore modeling

The training set consisting of 56 compounds was used to generate pharmacophore model. The HipHop(83) Algorithm available in Common Feature Pharmacophore Generation protocol tries to generate pharmacophore hypothesis (models) with features common amongst the molecules of the training set. During pharmacophore generation a minimum of 1 and a maximum of 10 pharmacophore feature such as hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), and hydrophobic (HY) were included. All the parameters were set to default. The HipHop algorithm scores each configuration based on both the degree to which it is common to the input set and its estimated rarity. The program begins by identifying configurations of features common to the molecules. More precisely, a configuration is a set of relative locations in 3D space, each associated with a type of feature. A molecule matches a configuration if it possesses a set of features and a conformation such that the set of features can be superimposed with the corresponding locations. A set of features is considered superimposed if each feature lies within a specified distance, the tolerance, from the corresponding ideal location. Certain molecules may be permitted to miss a feature until the total no. of molecules missing a feature does not drop below the set limit. The hypotheses are scored by classic maximum likelihood rule based on the probability of observed data under the assumption that the hypothesis is correct. The rarity of a configuration is estimated by applying regression model for perfect matches to the sub configuration including features in a subset. The algorithm can also optionally use information from inactive ligands to place excluded volume features. The HipHop algorithm reports the 10 top-scoring hypotheses until there is no improvement in

the hypothesis score. The resulting 10 pharmacophore hypotheses were used for the later steps of screening.

3.4 Database Selection

The database selected was DrugBank(112). DrugBank is a freely available web-enabled database that combines detailed drug data with comprehensive drug-target and drug-action information. It was specifically designed to facilitate in silico drug-target discovery, drug design, drug-metabolism prediction, drug-interaction prediction, and general pharmaceutical education. The database contains 6264 molecules drug entries including 1465 FDA-approved small molecule drugs, 132 FDA-approved biotech (protein/peptide) drugs, 86 nutraceuticals and 5076 experimental drugs. Since the screening is to be carried out to aid in the discovery of drugs for Alzheimer's disease, DrugBank was chosen as all the molecules will exhibit drug like property. All the available drug structures were used for the in-silico screening study. The structures of the drugs were downloaded in .sdf (Structural data file) format from the web database for use in the screening.

3.5 Local Database Creation

Database was created on the local workstation using the Build 3D database protocol in Discovery Studio 2.5. 6264 molecules were input for database creation. The database building was done with default parameters, except the conformation method was changed to FAST to get quick diverse low-energy conformations. During the building process 104 molecules showed error related to valency requirement or parsing error, hence were not included in the database created. Thus the final database created consisted of 6160 molecules.

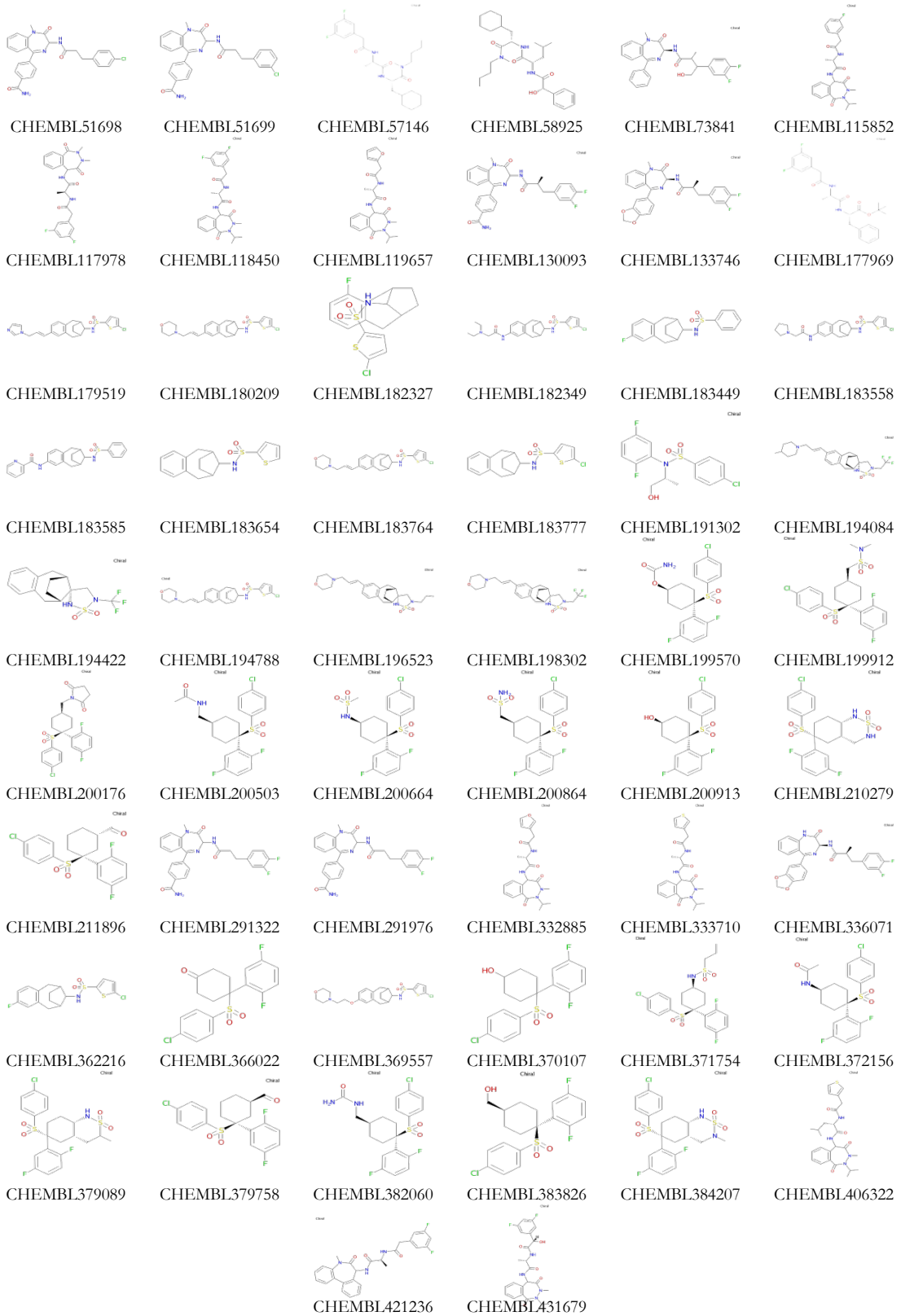

Figure 1 - **Structure of compounds in training set**. The two-dimensional chemical structure of 56 compounds taken in the training set. The ChEMBL Id for the compounds has been used to denote the compounds.

Table 1 - Training dataset for pharmacophore modelling for potential inhibitors of gamma-secretase showing experimental activity (IC50) and the total no. of pharmacophore features (Note - Exp.-Experimental, P.-Pharmacophore)

Compound Id	Exp. IC50 (nM)	No. of P. features
CHEMBL51698	15	14
CHEMBL51699	93	14
CHEMBL57146	32	8
CHEMBL58925	1	15
CHEMBL73841	6.9	13
CHEMBL115852	57	11
CHEMBL117978	5	11
CHEMBL118450	47	11
CHEMBL119657	86	11
CHEMBL130093	0.49	14
CHEMBL133746	13.2	13
CHEMBL177969	50	10
CHEMBL179519	15	20
CHEMBL180209	1	19
CHEMBL182327	34	16
CHEMBL182349	41	19
CHEMBL183449	70	12
CHEMBL183558	7	20
CHEMBL183585	75	18
CHEMBL183654	50	15
CHEMBL183764	12	20
CHEMBL183777	62	16
CHEMBL191302	27	14
CHEMBL194084	18	19
CHEMBL194422	17	16
CHEMBL194788	1	20
CHEMBL196523	19	19
CHEMBL198302	3.4	19

Compound Id	Exp. IC50 (nM)	No. of P. features
CHEMBL199570	1.9	11
CHEMBL199912	1.1	13
CHEMBL200176	1	10
CHEMBL200503	1.3	10
CHEMBL200664	4.2	20
CHEMBL200864	4.2	26
CHEMBL200913	10.7	19
CHEMBL210279	6.18	28
CHEMBL211896	55	10
CHEMBL291322	29	13
CHEMBL291976	71	26
CHEMBL332885	38	11
CHEMBL333710	14	11
CHEMBL336071	1.8	15
CHEMBL362216	29	16
CHEMBL366022	22	10
CHEMBL369557	5	20
CHEMBL370107	10	19
CHEMBL371754	0.82	18
CHEMBL372156	35	9
CHEMBL379089	0.08	26
CHEMBL379758	77	10
CHEMBL382060	1	12
CHEMBL383826	4.7	18
CHEMBL384207	0.9	24
CHEMBL406322	4	10
CHEMBL421236	76	13
CHEMBL431679	5	17

3.6 Database Screening

Instead of selecting a single pharmacophore hypothesis from the 10 hypotheses generated, all the pharmacophore hypotheses were used for screening the database created from the molecules obtained from DrugBank. The purpose of this screening is to retrieve hits for further screening. The database screening was carried out using Search 3D database protocol of Catalyst in Discovery Studio 2.5 with the query as the individual pharmacophore hypotheses; with default parameters except the search method was changed to FAST allowing a rigid fit of the ligand conformation against the pharmacophore. Thus 10 lists of compounds each corresponding to a pharmacophore could be obtained. Then a list with unique entires and entries with a Fit Value > 3.00 was prepared by merging all the lists of compounds obtained by screening DrugBank compounds with the pharmacophore as query.

3.7 Activity prediction using QSAR

The compounds in the training set were selected based on the available activity data i.e., experimental IC50 values. IC50 is the half maximal inhibitory concentration. It is a measure of effectiveness of a compound in inhibiting a biological function. It quantitatively gives the concentration of the drug or inhibitor required for the inhibition of a given biological process by half. FDA terms IC50 as the concentration of the drug required for 50% inhibition in vitro. IC50 can be related to affinity for competitive agonists and antagonists even though it is not direct indicator of affinity. The activity of the compounds (IC50) in the training set was for the production of Aβ-42 by cleaving of APP by gamma-secretase. Since the mechanism of action of gamma-secretase is known, the compound inhibits this biological process. The IC50 for the screened compound can be predicted by developing a QSAR model with the activity data of the training set of pharmacophore modeling as training set for QSAR model development. It is understood

that the development of a QSAR model which is a linear mathematical equation relating the biological activity of interest with the molecular descriptors of the compounds in the test set (unknown).

3.8 Calculation of Molecular descriptors

Molecular descriptors are required to be calculated for both the training set and the test set. The molecular descriptors were calculated for the compounds using e-Dragon tool. This tool can calculate over 3000 molecular descriptors which is lesser compared to the standalone dragon software. It required the input of the structure of the compound in SMILES or .sdf. The molecular descriptors were given as tab separated values. These were calculated and documented for both the test set and the training set. Some of the molecular descriptors calculated include number of rings (nCIC), number of rotable bonds (RBN), mean atomic polarizability (Mp), etc.

3.9 Selection of Molecular descriptors

The following molecular descriptors were selected from the list of over 3000 descriptors. Sum of atomic van der Waals volumes (Sv), mean atomic van der Waals volume (Mv), mean atomic Sanderson electronegativity (Me), mean atomic polarizability (Mp), number of bonds (nBT), number of rings (nCIC), hydrophilic factor (Hy), Moriguchi octanol-water partition coefficient (MLogP) and Ghose-Crippen octanol-water partition coefficient (ALogP). Thus a total of nine molecular properties were selected. The no. of molecular properties is one-fifth of the number of compounds in the training set. This is because multiple linear regression is employed to develop the QSAR model to predict the biological activity of the test set of compounds. The reason behind this could be understood from the following theory on multiple linear regression.

3.10 Multiple Linear Regression

Multiple linear regression is the most commonly employed method for constructing QSAR models. It is also a simple method. A multiple linear regression model is constructed with the assumption that a linear relationship exists between a set of molecular descriptors of a compound and a specific activity. The molecular descriptors are represented by a feature vector x with each descriptor as its component while the activity is represented by a quantity y. The following equation is a description of multiple linear regression model.

$$y = \alpha 0 + \alpha 1X1 + \alpha 2X2 + \cdots .. + \alpha nXn$$

where $\alpha 0$ is the regression model constant, ($X1, X2, \ldots.., Xn$) are molecular descriptors, $\alpha 1$ to $\alpha 2$ are the coefficients for the individual descriptors from $X1$ to Xn . The values for $\alpha 0$ to αn are chosen by minimizing the sum of squares of the residuals between the observed and predicted values defined by the equation so as to give the best prediction of y from x. This method has the advantage of having a simplistic form and being an easy interpretable expression. The positive or negative contribution of the molecular descriptors from $\alpha 1$ to αn is indicated by the sign of their respective coefficients. The relative importance of every descriptor to that activity is given by its magnitude. However, multiple linear regression works well only when the structure-activity relationship is linear in nature, the set of molecular descriptors are mathematically independent (orthogonal) of each another and the number of compounds in the training set exceeds the number of molecular descriptors by at least a factor of five. It has been found that, when collinear descriptors are used, the derived coefficients $\alpha 1$ to αn tend to be larger than the real values and sometimes have opposite signs. Therefore, the assumption of a linear relationship between a set of molecular descriptors and a specific activity may not always be appropriate, especially in the cases involving multiple mechanisms.

3.11 Developing QSAR model to predict IC50

QSAR model was developed by using the tool EasyQSAR v1.0. It uses the statistics engine of Microsoft Excel to perform the regression analysis. This requires the input of experimental biological data. In order to keep the mean deviation in the data of biological data as small as possible log IC50 is given as biological activity. The values of the molecular descriptors are fed in the tool. After performing the regression analysis the QSAR model that is a linear mathematical equation is given.

The QSAR model was then tested by substituting the test set molecular descriptor data in it. This was be used to validate the QSAR model by comparing the predicted and experimental activity data. Then the IC50 values were be predicted for the test set by substituting the values for the molecular descriptors in the QSAR model. The predicted IC50 values were used to filter the compounds for further screening.

3.12 Docking study with SwissDock

Docking study was carried out to find the binding of the ligand to the receptor. The challenge to carry out the study was the absence of complete three-dimensional structure of gamma-secretase complex. Sourcing of databases revealed the availability of structural data for the C-terminal fragment (CTF) of presenilin [PDB ID: 2KR6]. It had been identified that the catalytic core of the gamma-secretase complex is presenilin. The catalytic site of presenilin includes two aspartyl residues Asp 257 and Asp 285 on the 6^{th} and 7^{th} transmembrane domains respectively. The Asp 285 is present in the CTF of presenilin. Thus a compound that binds to Asp 285 or in the vicinity of Asp 285, block the binding of the substrate making it geometrically impossible. This involved the identification of binding site that possess Asp 285 using QsiteFinder tool(113). The geometric co-ordinates of the binding site with Asp 285 were predicted by the tool.

The compounds obtained after the filtering based on predicted IC50 values were used in docking study. The docking study was carried out using SwissDock(114). SwissDock is an EADock DSS(115) based web server. The algorithm consists of the following steps generation of binding modes wither in box or in the vicinity of all target cavities, estimation of CHARMM energies on a grid, evaluation of binding modes with most favorable energies and clustered, visualization of the favorable clusters. It allows the docking of protein receptor with ligand molecules. It requires the input format of the receptor to be .pdb (Protein Data Bankk) and that or ligand to be .mol2. These file format conversion were done using OpenBabel 2.3.1. The ligand and receptor files, minimized energetically, in the appropriate format were uploaded to the web-based server for docking study. The results contained the estimated free energy of binding (ΔG) and Full fitness score, besides the conformation of the ligand docked to the receptor. The lower estimated free energy of binding indicates the higher binding affinity. The compounds were again screened based on the estimated free energy of binding. The screening hits are suggested as potential inhibitors of gamma-secretase.

4 Results

4.1 Pharmacophore Modeling

The resulting pharmacophore hypotheses generated using the training set of active gamma-secretase inhibitors had a max fit score of 4 with rank ranging from 249.565 to 256.150 with the pharmacophore features hydrophobic center (HY) and hydrogen bond acceptor (HBA) and a total of 4 features. The configuration of the pharmacophore hypotheses are presented in (Figure 2) while (Table 2) gives the details of each hypothesis.

Table 2 - Pharmacophore hypotheses generated with the features present in the pharmacophore model and the maximum fit score along with the rank of each hypotheses.

Pharmacophore hypotheses	Features	Rank	Max Fit
01	HY HY HBA HBA	256.150	4
02	HY HY HBA HBA	253.297	4
03	HY HY HBA HBA	252.572	4
04	HY HY HBA HBA	252.518	4
05	HY HY HBA HBA	251.107	4
06	HY HY HBA HBA	251.098	4
07	HY HY HBA HBA	251.004	4
08	HY HY HBA HBA	250.902	4
09	HY HY HBA HBA	250.813	4
10	HY HY HBA HBA	249.565	4

HY – Hydrophobic center, HBA – Hydrogen Bond Acceptor

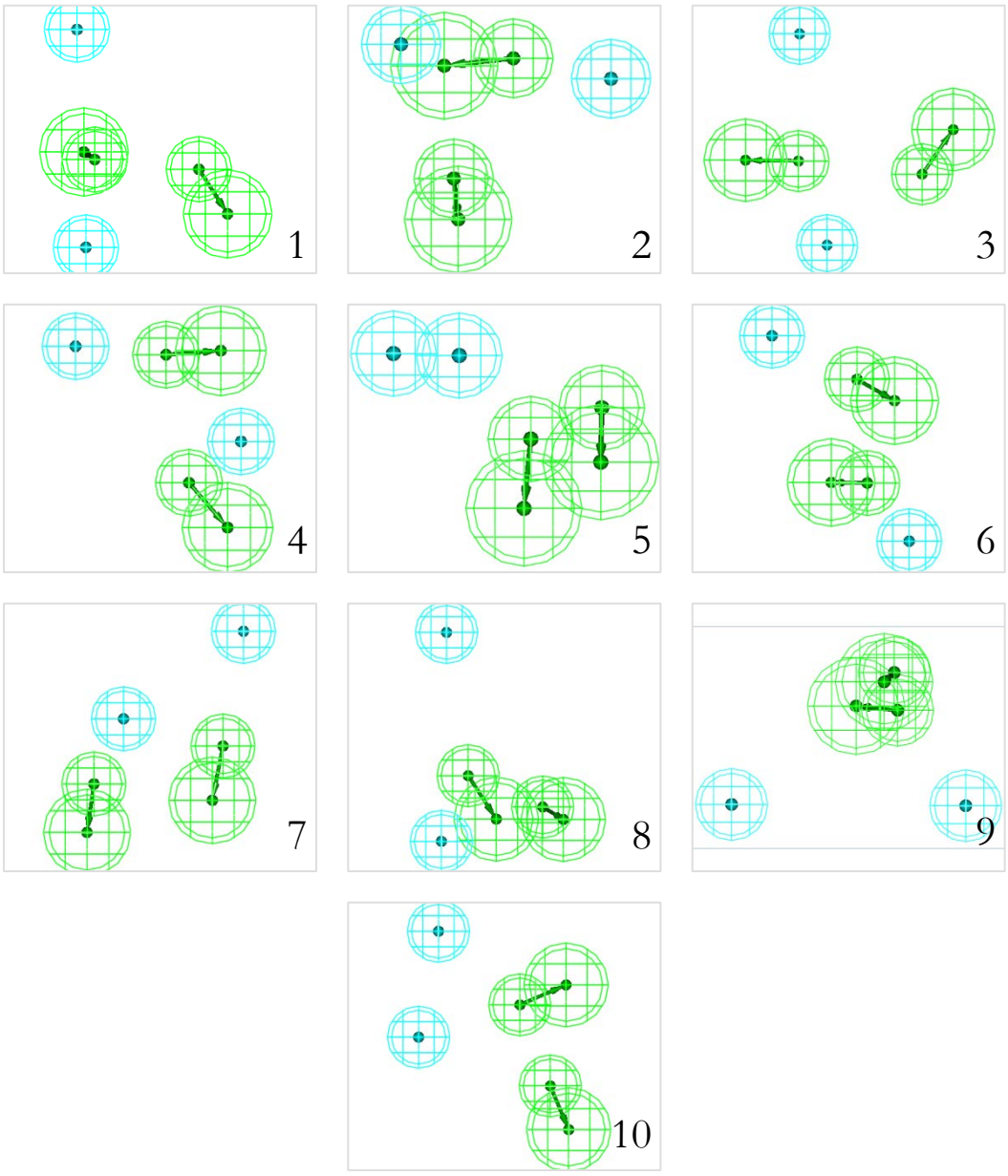

Figure 2 - **Pharmacophore hypotheses generated**. The input of 56 training set molecules gave 10 pharmacophore hypotheses. The pharmacophore hypotheses have hydrophobic, HY (cyan) and hydrogen bond acceptor, HBA (green). The total no. of features in each hypothesis is 4 (2 HY, 2 HBA).

4.2 Database Screening

4.2.1 Pharmacophore search

The screening of 6160 compounds from DrugBank with the pharmacophore hypotheses as query resulted in a reduced number of molecules than in total number of molecules in the database. No one hypotheses could be used alone as all the models had the same type of and the same number of pharmacophore features. Each hypothesis namely hypothesis 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 produced 1270, 1206, 1431, 1373, 1696, 1372, 1682, 1569, 1433 and 1832 hits in the screening process. The merging of the compound lists into another list with only compounds with fit value above 3.00 and unique entries contained 721 compounds. This list along with the fit value of the molecules is given in (Table 3)

4.3 Activity prediction using QSAR

The training set consisting of 54 compounds with experimental activity data (IC50) and molecular descriptor data was used to develop QSAR model. The training set compounds along with their experimental activity and with molecular descriptor data is given in (Table 4). EasyQSAR tool yielded a QSAR model upon input of data and followed by analysis. The following model was developed and can be used to predict IC50 of the 721 compounds in the test set. The QSAR model was validated by substituting the values of the training set in it. Thus the difference in experimental and predicted data was obtained. The values are given in (Table 6). The molecular descriptors of the compounds in the test set were calculated, and used to predict the IC50 value using the QSAR model. Their predicted IC50 values are given in (Table 7). The predicted IC50 ranged from 3.81E-08 to 3.60E+118. The IC50 prediction of DB00014 yielded no result due to erroneous values for molecular descriptors.

Table 3 – List of compounds screened from DrugBank using the pharmacophore hypotheses with Fit value > 3.0000 while 4.0000 is the maximum fit value

DrugBank Id	Fit Value
DB00014	3.27232
DB00175	3.46496
DB00177	3.19198
DB00178	3.33654
DB00183	3.43056
DB00197	3.28493
DB00203	3.10751
DB00206	3.64781
DB00212	3.7264
DB00213	3.18688
DB00214	3.46049
DB00220	3.49816
DB00222	3.51832
DB00224	3.40386
DB00225	3.18938
DB00227	3.43461
DB00243	3.30593
DB00251	3.16351
DB00261	3.51503
DB00274	3.28047
DB00278	3.2984
DB00287	3.46781
DB00309	3.0552
DB00317	3.42308
DB00320	3.23793
DB00337	3.14506
DB00342	3.30076
DB00346	3.5007
DB00364	3.13604
DB00372	3.00899
DB00385	3.59763
DB00391	3.1136
DB00418	3.53516
DB00438	3.39788
DB00439	3.43932
DB00444	3.72331
DB00445	3.52043
DB00447	3.13514
DB00451	3.03649
DB00457	3.61199
DB00471	3.14337
DB00481	3.28989
DB00492	3.41804
DB00502	3.53637
DB00527	3.10779
DB00528	3.48444
DB00541	3.02446
DB00549	3.35503
DB00559	3.43289
DB00562	3.13831
DB00587	3.08904
DB00590	3.45275
DB00596	3.71368
DB00607	3.12751
DB00619	3.36395
DB00622	3.51747
DB00644	3.2535
DB00646	3.45604
DB00661	3.04429
DB00680	3.02602
DB00682	3.24277
DB00688	3.21531
DB00691	3.29817
DB00694	3.53351
DB00696	3.11765
DB00705	3.32843
DB00718	3.16567
DB00732	3.31856
DB00762	3.36274
DB00767	3.18482
DB00770	3.61628
DB00773	3.81884
DB00775	3.03273
DB00796	3.01736
DB00802	3.13447
DB00803	3.45996
DB00807	3.38792
DB00826	3.32063
DB00836	3.62363
DB00862	3.36219
DB00881	3.23329
DB00882	3.57965
DB00901	3.41364
DB00905	3.42762
DB00910	3.29055
DB00912	3.31178
DB00914	3.1722
DB00929	3.15192
DB00932	3.64571
DB00938	3.20881
DB00948	3.47897
DB00950	3.41225
DB00973	3.03287
DB00976	3.16518
DB00983	3.29732
DB00997	3.60916
DB01016	3.12595
DB01029	3.3036
DB01039	3.60824
DB01046	3.38459
DB01067	3.00265
DB01070	3.45902
DB01072	3.32943
DB01083	3.5145
DB01088	3.30871
DB01089	3.42489
DB01095	3.15418
DB01098	3.17076
DB01101	3.35517
DB01118	3.36213
DB01128	3.07768
DB01130	3.43601
DB01134	3.02213
DB01135	3.15511
DB01136	3.62998
DB01138	3.30599
DB01148	3.28163
DB01149	3.22884
DB01160	3.38237
DB01162	3.5385
DB01167	3.06741
DB01177	3.63591
DB01180	3.24226
DB01200	3.01903
DB01226	3.2906
DB01248	3.34789
DB01249	3.66409
DB01251	3.53834
DB01254	3.11323
DB01256	3.44962
DB01258	3.89234
DB01263	3.26096
DB01264	3.28708
DB01274	3.20436
DB01287	3.50898
DB01295	3.22549
DB01319	3.25027
DB01321	3.22061
DB01324	3.25065
DB01326	3.58794
DB01329	3.63602
DB01332	3.2691
DB01340	3.73321
DB01342	3.35525
DB01347	3.6452
DB01348	3.09485
DB01349	3.41247
DB01361	3.07902
DB01362	3.00244
DB01388	3.3253
DB01396	3.29558
DB01410	3.19545
DB01411	3.15169
DB01416	3.33816
DB01419	3.49755
DB01425	3.52443
DB01459	3.22367
DB01462	3.29154
DB01464	3.37844
DB01487	3.45446
DB01505	3.19056

DB01523	3.04914
DB01570	3.25807
DB01578	3.04856
DB01601	3.50396
DB01602	3.61448
DB01604	3.13541
DB01612	3.77316
DB01627	3.52113
DB01635	3.15009
DB01660	3.41923
DB01683	3.27805
DB01691	3.2031
DB01701	3.09992
DB01709	3.26199
DB01712	3.30574
DB01713	3.23953
DB01716	3.40471
DB01737	3.62673
DB01738	3.46857
DB01745	3.20307
DB01751	3.36796
DB01755	3.80367
DB01762	3.5282
DB01812	3.2951
DB01819	3.1474
DB01829	3.24601
DB01839	3.10125
DB01842	3.03094
DB01858	3.29419
DB01864	3.40026
DB01871	3.36834
DB01872	3.02879
DB01883	3.27621
DB01891	3.46002
DB01892	3.24711
DB01914	3.28104
DB01920	3.64399
DB01922	3.12406
DB01940	3.60919
DB01965	3.0002
DB01979	3.27231
DB01991	3.15187
DB02008	3.10181
DB02017	3.18798
DB02034	3.3265
DB02041	3.10219
DB02046	3.29579
DB02056	3.1904
DB02057	3.59205
DB02062	3.12782
DB02065	3.30082
DB02071	3.0205
DB02082	3.63258
DB02095	3.39165
DB02104	3.02905
DB02115	3.20064
DB02139	3.09955
DB02140	3.2175
DB02151	3.39826
DB02169	3.33178
DB02177	3.01818
DB02210	3.16211
DB02226	3.23326
DB02229	3.21737
DB02236	3.64682
DB02237	3.11652
DB02253	3.38631
DB02255	3.20443
DB02258	3.38961
DB02263	3.47201
DB02301	3.12812
DB02333	3.51231
DB02348	3.08521
DB02349	3.00334
DB02355	3.41628
DB02370	3.63866
DB02385	3.04737
DB02395	3.16159
DB02397	3.28321
DB02423	3.12834
DB02477	3.27328
DB02484	3.41277
DB02501	3.27513
DB02545	3.4691
DB02549	3.1614
DB02550	3.05121
DB02552	3.03335
DB02555	3.35905
DB02558	3.10085
DB02565	3.04595
DB02590	3.21659
DB02597	3.52131
DB02620	3.44382
DB02632	3.1861
DB02641	3.35132
DB02663	3.30838
DB02669	3.25089
DB02696	3.29821
DB02697	3.36541
DB02703	3.165
DB02706	3.45289
DB02717	3.76578
DB02721	3.10562
DB02733	3.17963
DB02777	3.3933
DB02783	3.06979
DB02787	3.1553
DB02791	3.13738
DB02804	3.22295
DB02826	3.08476
DB02848	3.49303
DB02855	3.1331
DB02858	3.06817
DB02884	3.02029
DB02892	3.13757
DB02894	3.38135
DB02907	3.29678
DB02922	3.16342
DB02932	3.07768
DB02937	3.1044
DB02946	3.443
DB02975	3.5582
DB02979	3.69251
DB02993	3.42749
DB03018	3.23333
DB03029	3.10897
DB03031	3.03559
DB03053	3.24762
DB03070	3.20834
DB03077	3.44258
DB03090	3.66922
DB03103	3.46639
DB03110	3.37712
DB03133	3.65706
DB03178	3.31986
DB03189	3.32561
DB03202	3.0995
DB03240	3.35352
DB03249	3.11133
DB03300	3.36748
DB03309	3.46019
DB03319	3.56888
DB03321	3.30974
DB03336	3.25273
DB03338	3.03826
DB03343	3.21154
DB03344	3.00026
DB03359	3.21412
DB03367	3.09658
DB03382	3.30717
DB03383	3.05495
DB03393	3.32487
DB03431	3.22121
DB03434	3.37266
DB03440	3.01773
DB03446	3.02565
DB03450	3.5249
DB03459	3.4437
DB03520	3.43233
DB03537	3.20394
DB03541	3.12176
DB03550	3.26701
DB03554	3.23361
DB03556	3.38043
DB03564	3.40158
DB03573	3.27897
DB03618	3.15675
DB03629	3.20871
DB03648	3.06799
DB03673	3.16483
DB03691	3.22832
DB03692	3.37437
DB03696	3.10655
DB03708	3.41382

DB03730	3.1459
DB03732	3.32688
DB03735	3.21408
DB03741	3.53202
DB03744	3.36611
DB03754	3.41443
DB03767	3.4588
DB03785	3.48979
DB03793	3.20469
DB03796	3.13682
DB03814	3.52181
DB03819	3.20061
DB03830	3.32126
DB03837	3.42291
DB03855	3.45094
DB03862	3.26126
DB03866	3.3285
DB03870	3.39905
DB03878	3.07585
DB03880	3.38237
DB03890	3.01436
DB03906	3.30576
DB03919	3.29621
DB03934	3.20265
DB03943	3.30424
DB03944	3.34285
DB03971	3.49458
DB03984	3.16249
DB03991	3.35361
DB03995	3.10207
DB04012	3.4654
DB04015	3.14791
DB04036	3.01542
DB04049	3.05244
DB04075	3.44713
DB04084	3.44516
DB04092	3.43712
DB04093	3.49894
DB04133	3.01139
DB04139	3.32542
DB04140	3.4541
DB04145	3.49873
DB04156	3.44504
DB04179	3.23004
DB04187	3.40518
DB04189	3.16964
DB04200	3.36304
DB04219	3.71055
DB04231	3.2196
DB04232	3.64386
DB04234	3.18414
DB04242	3.36753
DB04261	3.2799
DB04270	3.01639
DB04276	3.32066
DB04299	3.49912
DB04310	3.01513
DB04317	3.25154
DB04320	3.47032
DB04369	3.29803
DB04386	3.29139
DB04416	3.07406
DB04421	3.43362
DB04425	3.06762
DB04441	3.6035
DB04485	3.29038
DB04487	3.59815
DB04497	3.31768
DB04498	3.52406
DB04501	3.49731
DB04505	3.41539
DB04517	3.21327
DB04529	3.04636
DB04534	3.51577
DB04542	3.56679
DB04553	3.28683
DB04570	3.37709
DB04572	3.01147
DB04574	3.66919
DB04579	3.15399
DB04580	3.05703
DB04590	3.39385
DB04592	3.433
DB04594	3.51932
DB04595	3.59709
DB04609	3.45001
DB04619	3.27595
DB04622	3.32661
DB04624	3.19172
DB04626	3.42554
DB04632	3.36163
DB04642	3.34443
DB04643	3.14372
DB04644	3.08905
DB04645	3.6189
DB04657	3.36982
DB04658	3.09994
DB04673	3.43867
DB04674	3.42552
DB04680	3.23778
DB04693	3.15864
DB04695	3.25592
DB04732	3.42656
DB04738	3.35933
DB04739	3.12429
DB04746	3.38446
DB04751	3.27724
DB04771	3.12344
DB04773	3.38935
DB04781	3.44767
DB04789	3.31897
DB06699	3.20471
DB06717	3.26503
DB06829	3.14483
DB06835	3.44175
DB06844	3.27958
DB06846	3.10273
DB06859	3.01745
DB06886	3.31257
DB06892	3.01101
DB06909	3.24602
DB06953	3.06518
DB06959	3.31184
DB06960	3.37555
DB06961	3.09784
DB06964	3.13945
DB06969	3.29195
DB06970	3.1788
DB06974	3.53364
DB06979	3.0599
DB06986	3.01438
DB06993	3.06086
DB07041	3.25019
DB07046	3.01165
DB07056	3.5143
DB07062	3.26934
DB07066	3.24501
DB07068	3.14562
DB07089	3.26274
DB07090	3.54121
DB07093	3.30163
DB07105	3.19883
DB07111	3.11053
DB07113	3.34901
DB07160	3.42859
DB07165	3.17255
DB07169	3.04386
DB07172	3.00759
DB07177	3.27814
DB07188	3.21676
DB07209	3.13653
DB07219	3.35444
DB07221	3.1903
DB07223	3.38504
DB07224	3.3704
DB07225	3.36386
DB07231	3.38987
DB07246	3.13994
DB07249	3.12858
DB07250	3.38841
DB07251	3.07339
DB07255	3.09591
DB07264	3.50321
DB07269	3.42965
DB07272	3.24546
DB07276	3.28224
DB07278	3.35521
DB07309	3.27338
DB07321	3.2265
DB07325	3.03097
DB07333	3.55434
DB07334	3.43394
DB07341	3.43029
DB07349	3.21973

DB07351	3.53153
DB07359	3.34329
DB07362	3.13228
DB07369	3.13765
DB07370	3.05719
DB07377	3.16709
DB07379	3.16904
DB07394	3.13371
DB07397	3.11832
DB07403	3.46466
DB07413	3.31096
DB07416	3.45423
DB07429	3.02571
DB07432	3.61073
DB07446	3.32011
DB07448	3.2955
DB07460	3.0708
DB07471	3.26785
DB07475	3.40246
DB07520	3.22486
DB07537	3.42026
DB07541	3.36487
DB07558	3.34264
DB07571	3.22055
DB07572	3.69494
DB07574	3.02731
DB07582	3.29136
DB07583	3.31588
DB07602	3.45365
DB07608	3.03313
DB07615	3.00809
DB07624	3.16186
DB07627	3.10493
DB07630	3.38644
DB07633	3.21023
DB07640	3.11709
DB07657	3.33817
DB07659	3.45481
DB07675	3.30853
DB07691	3.43133
DB07692	3.01429
DB07713	3.11644
DB07724	3.38255
DB07735	3.4419
DB07738	3.5011
DB07740	3.07738
DB07750	3.15467
DB07760	3.08613
DB07765	3.38477
DB07769	3.01401
DB07770	3.12738
DB07771	3.0657
DB07772	3.09071
DB07778	3.4989
DB07779	3.48662
DB07780	3.44264
DB07784	3.01295
DB07786	3.14929
DB07792	3.57357
DB07793	3.59994
DB07800	3.24832
DB07801	3.48707
DB07804	3.08712
DB07826	3.05045
DB07827	3.05965
DB07831	3.30738
DB07832	3.55494
DB07836	3.31347
DB07839	3.33848
DB07841	3.47019
DB07844	3.18774
DB07845	3.10007
DB07846	3.15458
DB07847	3.09737
DB07848	3.3106
DB07853	3.15355
DB07861	3.08314
DB07870	3.01414
DB07875	3.03979
DB07878	3.15234
DB07881	3.53875
DB07895	3.34934
DB07898	3.5006
DB07916	3.53539
DB07926	3.33305
DB07934	3.09244
DB07936	3.6422
DB07943	3.29641
DB07946	3.19846
DB07947	3.30277
DB07962	3.04874
DB07964	3.17217
DB07967	3.3019
DB07970	3.23301
DB07974	3.32503
DB07981	3.68001
DB07982	3.0002
DB07984	3.49608
DB07988	3.08558
DB07995	3.02247
DB08001	3.07903
DB08005	3.3147
DB08009	3.50025
DB08013	3.42261
DB08014	3.32433
DB08026	3.2574
DB08078	3.17402
DB08081	3.19259
DB08086	3.27663
DB08090	3.26292
DB08105	3.17016
DB08116	3.58996
DB08118	3.20655
DB08143	3.37454
DB08147	3.28084
DB08152	3.00496
DB08154	3.47616
DB08163	3.01833
DB08169	3.3956
DB08170	3.39782
DB08173	3.11152
DB08174	3.17037
DB08177	3.17126
DB08180	3.32813
DB08185	3.39356
DB08195	3.37956
DB08221	3.22217
DB08223	3.48597
DB08226	3.37374
DB08229	3.11298
DB08240	3.15195
DB08266	3.3126
DB08275	3.13276
DB08278	3.00038
DB08282	3.02352
DB08302	3.02876
DB08310	3.05005
DB08322	3.49423
DB08324	3.06359
DB08325	3.11189
DB08341	3.05948
DB08342	3.27843
DB08344	3.43494
DB08349	3.37915
DB08350	3.11552
DB08354	3.03027
DB08366	3.39804
DB08375	3.26791
DB08376	3.36066
DB08379	3.13324
DB08387	3.32327
DB08388	3.28485
DB08403	3.10851
DB08404	3.56304
DB08407	3.17371
DB08420	3.29087
DB08429	3.17177
DB08435	3.30676
DB08437	3.31201
DB08438	3.06511
DB08444	3.5952
DB08446	3.42654
DB08457	3.46551
DB08458	3.1417
DB08459	3.54916
DB08460	3.59298
DB08463	3.0894
DB08464	3.25917
DB08465	3.50273
DB08471	3.304
DB08477	3.17603
DB08481	3.39578
DB08482	3.22652
DB08483	3.15879

DB08487	3.15312
DB08489	3.17874
DB08490	3.03529
DB08499	3.27014
DB08502	3.00108
DB08505	3.07571
DB08526	3.35635
DB08556	3.2513
DB08582	3.08876
DB08585	3.11964
DB08586	3.11073
DB08590	3.68935
DB08614	3.05093

DB08622	3.41244
DB08633	3.34726
DB08640	3.41984
DB08643	3.29804
DB08664	3.14168
DB08669	3.13622
DB08670	3.60572
DB08690	3.10672
DB08704	3.54552
DB08706	3.08596
DB08710	3.19863
DB08719	3.01329
DB08720	3.02005

DB08721	3.42443
DB08722	3.26748
DB08727	3.2072
DB08732	3.09532
DB08733	3.69927
DB08749	3.5617
DB08755	3.47593
DB08762	3.36458
DB08772	3.71016
DB08775	3.12828
DB08810	3.64242
DB08811	3.52842

Table 4 – Training Set compounds of QSAR model development with activity (IC50) and molecular descriptor data (Sv, Mv, Me, Mp, nBt, nCIC, Hy, MLogP, ALogP)

Compound Id	log IC50	Sv	Mv	Me	Mp	nBT	nCIC	Hy	MLogP	ALogP
CHEMBL51698	1.176091	38.19	0.67	1.01	0.7	60	4	0.811	3.668	3.934
CHEMBL51699	1.968483	38.19	0.67	1.01	0.7	60	4	0.811	3.668	3.934
CHEMBL57146	1.50515	40.5	0.58	1	0.61	71	2	0.192	3.398	4.549
CHEMBL58925	0	45.58	0.57	0.99	0.61	81	2	0.763	2.503	4.824
CHEMBL73841	0.838849	38.91	0.65	1.02	0.67	63	4	0.162	4.289	4.235
CHEMBL115852	1.755875	37.31	0.62	1.01	0.65	62	3	0.222	2.494	1.956
CHEMBL117978	0.69897	34.22	0.63	1.03	0.65	56	3	0.27	2.447	1.435
CHEMBL118450	1.672098	37.42	0.62	1.02	0.65	62	3	0.235	2.866	2.162
CHEMBL119657	1.934498	35.11	0.62	1.01	0.64	59	3	0.257	1.031	1.146
CHEMBL130093	-0.3098	39.31	0.66	1.02	0.68	63	4	0.801	3.866	4.144
CHEMBL133746	1.120574	38.83	0.66	1.02	0.68	63	5	-0.355	3.936	4.909
CHEMBL177969	1.69897	36.63	0.61	1.02	0.64	61	2	0.208	3.759	3.92
CHEMBL179519	1.176091	36.46	0.66	1	0.72	59	5	-0.329	3.857	4.513
CHEMBL180209	0	38.77	0.64	1	0.69	65	5	-0.341	3.403	4.598
CHEMBL182327	1.531479	27.39	0.67	1.02	0.73	44	4	-0.26	3.925	4.574
CHEMBL182349	1.612784	38.76	0.63	1	0.68	65	4	0.239	3.097	4.406
CHEMBL183449	1.845098	28.19	0.64	1	0.68	47	4	-0.342	3.809	4.329
CHEMBL183558	0.845098	38.17	0.64	1	0.69	64	5	0.239	3.097	4.17
CHEMBL183585	1.875061	37.18	0.65	1	0.69	61	5	0.177	3.011	4.186
CHEMBL183654	1.69897	26.57	0.65	1	0.71	44	4	-0.31	3.035	3.91
CHEMBL183764	1.079181	36.37	0.67	1.01	0.72	58	5	0.239	3.143	4.431
CHEMBL183777	1.792392	27.28	0.67	1	0.73	44	4	-0.284	3.54	4.369
CHEMBL191302	1.431364	24.32	0.66	1.04	0.7	38	2	-0.198	3.345	4.057
CHEMBL194084	1.255273	40.59	0.6	1.01	0.64	72	5	-0.334	4.547	5.055
CHEMBL194422	1.230449	24.81	0.62	1.04	0.66	43	4	-0.198	3.218	4.473
CHEMBL194788	0	38.77	0.64	1	0.69	65	5	-0.341	3.403	4.598
CHEMBL196523	1.278754	39.17	0.59	0.99	0.64	70	5	-0.361	3.262	3.438
CHEMBL198302	0.531479	37.9	0.6	1.02	0.64	67	5	-0.292	3.367	3.573
CHEMBL199570	0.278754	30.03	0.65	1.04	0.69	48	3	0.318	3.823	4.662
CHEMBL199912	0.041393	34.91	0.63	1.03	0.69	57	3	-0.671	4.182	5.076
CHEMBL200176	0	35.23	0.65	1.03	0.69	57	4	-0.712	4.601	4.595
CHEMBL200503	0.113943	32.71	0.64	1.02	0.68	53	3	-0.302	4.638	4.555
CHEMBL200664	0.623249	31.72	0.65	1.03	0.7	51	3	-0.234	3.749	4.21
CHEMBL200864	0.623249	31.72	0.65	1.03	0.7	51	3	0.33	3.749	4.263
CHEMBL200913	1.029384	27.52	0.66	1.03	0.7	44	3	-0.277	4.459	4.493
CHEMBL210279	0.790988	32.11	0.66	1.04	0.7	52	4	0.342	4.195	3.246
CHEMBL211896	1.740363	28.52	0.66	1.03	0.71	45	3	-0.722	4.606	4.936
CHEMBL291322	1.462398	37.71	0.66	1.02	0.68	60	4	0.82	3.668	3.681
CHEMBL291976	1.851258	34.81	0.61	1	0.64	59	3	1.595	2.068	2.586
CHEMBL332885	1.579784	35.11	0.62	1.01	0.64	59	3	0.257	1.031	0.854
CHEMBL333710	0.255273	37.23	0.66	1.03	0.68	60	5	0.191	3.738	4.704
CHEMBL336071	1.462398	27.39	0.67	1.02	0.73	44	4	-0.26	3.925	4.574
CHEMBL362216	0.69897	38.28	0.63	1.01	0.68	65	5	-0.31	2.768	4.179
CHEMBL366022	1	27.52	0.66	1.03	0.7	44	3	-0.277	4.459	4.493
CHEMBL369557	-0.08619	34.91	0.63	1.03	0.69	57	3	-0.265	4.182	5.082
CHEMBL370107	1.544068	31.12	0.65	1.03	0.69	50	3	-0.288	4.423	4.224
CHEMBL371754	-1.09691	34.31	0.65	1.03	0.7	56	4	-0.265	4.182	4.495
CHEMBL372156	1.886491	28.52	0.66	1.03	0.71	45	3	-0.722	4.606	4.936
CHEMBL379089	0	32.11	0.64	1.03	0.68	52	3	0.96	4.041	4.345
CHEMBL379758	0.672098	29.12	0.65	1.03	0.69	47	3	-0.294	4.682	4.823
CHEMBL382060	-0.04576	33.71	0.65	1.03	0.7	55	4	-0.233	4.413	3.782
CHEMBL383826	0.60206	40.48	0.61	1	0.65	68	3	0.206	2.403	2.644
CHEMBL384207	1.880814	37.31	0.65	1.02	0.68	60	4	0.177	3.417	3.491
CHEMBL406322	0.69897	37.93	0.62	1.03	0.64	63	3	0.877	2.127	1.616

QSAR model to predict IC50 value of potential inhibitors of gamma secretase from DrugBank

logIC50 = 5.333468493034E+001 + 9.441980133615E–002***Sv**

+ 1.615431581663E+001***Mv** + – 4.127786888796E+001***Me**

+ –2.451283848884E+001***Mp** + –1.253675687200E–001***nBT**

+ 3.276417714217E–002***nCIC** + –1.388896872996E–001***Hy**

+ 4.968101794863E–002***MLOGP** + –1.111794822900E–001***ALOGP**

The following table serves to give a brief account of the various terms in the QSAR model developed to predict IC50.

Table 5 - Molecular descriptors used in QSAR model along with its coefficient and input data range

Descriptor	Explanation	Range	Coefficient
Sv	sum of atomic van der Waals volumes	24.320 – 45.580	9.441980133615E–002
Mv	mean atomic van der Waals volume	00.570 – 00.670	1.615431581663E+001
Me	mean atomic Sanderson electronegativity	00.990 – 01.040	–4.127786888796E+001
Mp	mean atomic polarizability	00.610 – 00.730	–2.451283848884E+001
nBT	number of bonds	38.000 – 81.000	–1.253675687200E–001
nCIC	number of rings	02.000 – 05.000	3.276417714217E–002
Hy	hydrophilic factor	-00.722 – 01.595	–1.388896872996E–001
MLOGP	Moriguchi octanol-water partition coeffecient	01.031 – 04.682	4.968101794863E–002
ALOGP	Ghose-Crippen octanol-water partition coeffecient	00.854 – 05.082	–1.111794822900E–001

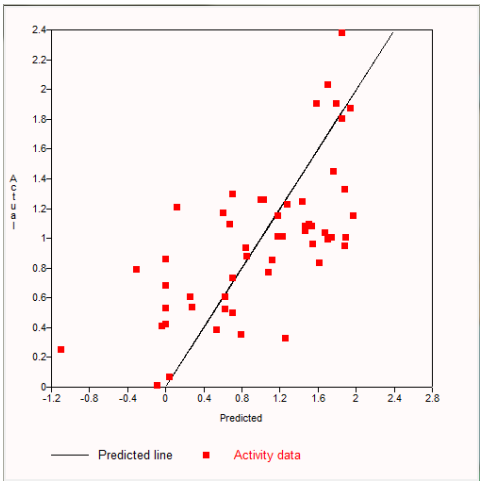

Figure 3 - Plot of Actual activity data (log IC50) vs Predicted activity data (log IC50)

Table 6 - Validation of QSAR model. Actual activity data (log IC50) and Predicted Activity data (log IC50)

Compound Id	Actual IC50	Predicted IC50
CHEMBL51698	1.18	1.16
CHEMBL51699	1.97	1.16
CHEMBL57146	1.51	1.10
CHEMBL58925	0.00	0.42
CHEMBL73841	0.84	0.93
CHEMBL115852	1.76	1.45
CHEMBL117978	0.70	1.30
CHEMBL118450	1.67	1.04
CHEMBL119657	1.93	1.88
CHEMBL130093	-0.31	0.79
CHEMBL133746	1.12	0.86
CHEMBL177969	1.70	1.00
CHEMBL179519	1.18	1.01
CHEMBL180209	0.00	0.86
CHEMBL182327	1.53	1.08
CHEMBL182349	1.61	0.84
CHEMBL183449	1.85	2.38
CHEMBL183558	0.85	0.88
CHEMBL183585	1.88	1.33
CHEMBL183654	1.70	2.04
CHEMBL183764	1.08	0.78
CHEMBL183777	1.79	1.91
CHEMBL191302	1.43	1.25
CHEMBL194084	1.26	0.33
CHEMBL194422	1.23	1.02
CHEMBL194788	0.00	0.86
CHEMBL196523	1.28	1.23
CHEMBL198302	0.53	0.39
CHEMBL199570	0.28	0.54
CHEMBL199912	0.04	0.07
CHEMBL200176	0.00	0.53
CHEMBL200503	0.11	1.21
CHEMBL200664	0.62	0.61
CHEMBL200864	0.62	0.53
CHEMBL200913	1.03	1.26
CHEMBL210279	0.79	0.35
CHEMBL211896	1.74	1.01
CHEMBL291322	1.46	1.05
CHEMBL291976	1.85	1.80
CHEMBL332885	1.58	1.91
CHEMBL333710	0.26	0.60
CHEMBL336071	1.46	1.08
CHEMBL362216	0.70	0.50
CHEMBL366022	1.00	1.26
CHEMBL369557	-0.09	0.01
CHEMBL370107	1.54	0.96
CHEMBL371754	-1.10	0.25
CHEMBL372156	1.89	1.01
CHEMBL379089	0.00	0.68
CHEMBL379758	0.67	1.10
CHEMBL382060	-0.05	0.41
CHEMBL383826	0.60	1.17
CHEMBL384207	1.88	0.95
CHEMBL406322	0.70	0.73

Table 7 – Predicted IC50 values of the hits of pharmacophore screen of DrugBank compounds

DrugBank Id	Predicted IC50
DB02348	3.81E-08
DB00732	1.14E-06
DB00976	0.000311711
DB00337	0.000535658
DB01321	0.001612038
DB02777	0.004394264
DB01167	0.005202091
DB00541	0.005233743
DB01751	0.009622836
DB01263	0.015164695
DB01180	0.026539552
DB02922	0.036372022
DB00206	0.054918572
DB00492	0.077143903
DB04674	0.090207227
DB01410	0.090997384
DB01419	0.097180046
DB01089	0.114075128
DB02169	0.128672916
DB00528	0.130116387
DB01892	0.131788053
DB01083	0.146228089
DB07692	0.169452007
DB07169	0.187343009
DB01256	0.203074233
DB00251	0.208713409
DB01118	0.235493837
DB00762	0.244249327
DB02423	0.275189756
DB02975	0.329850735
DB00932	0.386193534
DB02545	0.469168469
DB01388	0.577240369
DB01459	0.662676154
DB07321	0.716987422
DB03031	0.730506228
DB01396	0.75426413
DB00444	0.755936862
DB03696	0.897846467
DB00661	0.916346672
DB00471	0.926668223
DB01130	0.956238307
DB00549	1.061458814
DB08266	1.111649142
DB01134	1.114938101
DB06717	1.223170661
DB08195	1.550482562
DB01149	1.599870221
DB01248	1.813602601
DB00862	1.931435479
DB06953	2.038262796
DB07870	2.137557657
DB01712	2.237392688
DB08772	2.305665711
DB08721	2.450858749
DB00559	2.52673191
DB00203	2.545069012
DB07847	2.681155896
DB00773	2.91742236
DB07394	2.971823372
DB00213	3.08640754
DB07249	3.176551166
DB04673	3.258753317
DB01070	3.366556309
DB01148	3.603814182
DB08446	3.673846444
DB01347	3.690095726
DB08177	3.786500964
DB07369	3.923531979
DB02177	4.008961192
DB07278	4.015508643
DB07359	4.036080707
DB00596	4.12517379
DB07844	4.150280444
DB08013	4.214840332
DB00767	4.232942434
DB00802	4.37897904
DB08477	4.387232583
DB00836	4.433389531
DB00317	4.462990884
DB08459	4.782383387
DB07675	4.801973912
DB08722	4.803108446
DB00696	4.835972695
DB08483	4.90981081
DB01251	4.923074914
DB07093	5.027065633
DB03837	5.232905149
DB02721	5.383614999
DB01200	5.424874924
DB08344	5.589559992
DB02139	5.671880056
DB00342	5.755347538
DB08457	5.761958253
DB00372	5.834282601
DB08811	6.142314866
DB06970	6.143585093
DB00451	6.151493595
DB04609	6.334569825
DB01039	6.339743437
DB00688	6.498383105
DB03520	6.522469786
DB00680	6.611450381
DB01462	6.858932275
DB08481	6.98256853
DB07334	7.050592035
DB00227	7.230824685
DB08387	7.598323948
DB00691	7.618034224
DB08154	7.62023424
DB00320	7.681857778
DB03944	7.72808924
DB08471	7.972992125
DB07403	8.154936986
DB01138	8.283567678
DB03336	8.374951775
DB07827	8.623147018
DB02804	8.837301517
DB01088	9.536999584
DB08014	10.18023054
DB07836	10.41213603
DB02703	10.60927423
DB01523	10.79536273
DB00882	10.85080715
DB07349	11.09156596
DB00481	11.24452046
DB06974	11.86151121
DB07219	11.89067029
DB00197	12.02332796
DB07724	12.37127699
DB07769	12.53901026
DB01348	13.04481577
DB01128	13.05211767
DB02932	13.05211767
DB00718	13.22966021
DB07800	13.7824937
DB03554	14.04092284
DB01324	14.15607988
DB00309	14.20273869
DB04738	14.24531646
DB04270	14.27527693
DB00385	14.44730607
DB01601	14.49037116
DB08556	14.68449556
DB08152	14.78550344
DB01464	14.80037714
DB08719	14.86436866
DB07471	15.04866161
DB00901	15.60140331
DB04232	15.67497266
DB08502	15.72455769
DB08582	16.09107524
DB08585	16.09107524
DB08633	16.67259795
DB03890	16.70290996
DB03556	16.71159972
DB07846	17.54329901
DB08143	17.54502312
DB07936	18.24600405
DB07056	19.07653729

DB07250	19.22667546
DB01072	19.28745327
DB07309	19.50751145
DB06892	20.22436475
DB00622	20.42667256
DB00274	20.43166373
DB02226	20.74486169
DB02550	22.0780343
DB08278	22.426288
DB07333	22.71186878
DB08727	23.19140953
DB08690	23.55496864
DB02253	23.91832438
DB07962	24.15394943
DB04574	24.24609981
DB01029	24.54498881
DB08458	25.17730444
DB00224	25.44304671
DB02791	25.95292883
DB06844	26.04699153
DB08282	26.40537153
DB08755	27.23316175
DB07804	27.58996761
DB00590	27.59213078
DB04595	27.62018058
DB01411	28.17649928
DB08221	28.72993131
DB02894	29.79939876
DB07974	30.32384298
DB08226	30.41643688
DB00287	30.81642505
DB04317	31.51791371
DB07255	31.60122012
DB00950	32.69915131
DB07640	32.78124856
DB08490	33.37865711
DB07691	34.00457028
DB02706	34.29146068
DB04739	34.79638272
DB00243	35.32849204
DB08444	35.33570142
DB00973	35.49738772
DB08349	36.83138578
DB01342	36.9897858
DB00587	37.15874359
DB07770	37.19607425
DB00912	38.61783112
DB07778	39.07279091
DB07967	39.95642035
DB08379	40.70629863
DB08147	40.87515853
DB01570	41.04681715
DB07826	41.12785243
DB07848	41.19177745
DB00948	41.72430157
DB07981	41.97247274
DB07984	41.97247274
DB01329	42.13299469
DB00619	43.91280048
DB02848	44.66188467
DB08310	45.74398685
DB07276	46.51766058
DB07541	46.74678646
DB04012	47.52140224
DB08376	47.9562035
DB02858	48.13094039
DB00222	48.13921289
DB07583	48.19615105
DB08720	49.05134228
DB02008	49.43699948
DB00502	49.52883072
DB01340	51.64215073
DB00562	52.96611132
DB07188	55.25928939
DB00607	56.7945971
DB04140	57.16452076
DB00212	57.49420562
DB08590	57.67619247
DB01691	57.74990297
DB03984	58.17947578
DB07062	58.91035754
DB08710	59.0263144
DB00261	60.43684855
DB04505	61.10735255
DB02071	63.60350033
DB07090	65.73976713
DB02946	66.57790175
DB04642	66.7405177
DB07429	67.12810397
DB01349	67.18304893
DB07574	68.00611705
DB00881	69.24225124
DB00220	71.47864857
DB07964	73.27199306
DB07251	74.46867861
DB00177	75.60274692
DB08174	78.34161895
DB07878	78.76157103
DB08350	78.77990504
DB08388	79.36483516
DB00929	79.57301041
DB08706	84.93159871
DB01016	86.58658764
DB04693	88.01733772
DB00178	89.06213468
DB00527	91.66379965
DB00796	91.84092942
DB07068	95.55264233
DB08614	95.71894318
DB08622	95.71894318
DB01046	97.41320828
DB08302	100.641473
DB08078	101.2015464
DB01264	105.9903972
DB04416	110.1857888
DB08173	111.7348277
DB08435	113.6742946
DB03866	119.3144066
DB07947	120.5958566
DB07624	122.7609298
DB07832	123.255345
DB07875	128.1088405
DB07582	130.3728379
DB07841	133.5920292
DB03367	136.1975616
DB00682	137.9838223
DB08341	139.9408802
DB08086	140.4423279
DB04643	142.880314
DB08749	144.4133736
DB02258	145.570467
DB02140	148.2422977
DB08487	152.7575766
DB00910	153.1173899
DB07223	154.2000716
DB07988	154.3862726
DB01871	154.9262242
DB06961	155.3522808
DB04644	155.412954
DB04645	156.6111216
DB08223	156.9200781
DB07165	157.0444257
DB06886	159.8542878
DB04234	160.2565418
DB04632	160.4122741
DB07982	161.9992285
DB00439	168.9949342
DB01098	169.677065
DB07416	170.3442087
DB00775	171.2566738
DB00705	179.2778369
DB03785	183.2491536
DB00457	183.722619
DB07362	186.4105304
DB07630	189.3272967
DB01254	198.27068
DB01505	198.4732809
DB04746	199.4601831
DB00438	208.6486923
DB07446	209.1618455
DB07943	209.8692913
DB03383	210.8040978
DB07995	215.4545106
DB08324	224.4440867
DB07839	226.3066801
DB08664	227.3635251
DB06829	227.6389824
DB08460	231.7184559
DB06846	232.6155366
DB08429	234.0071243
DB01162	234.5518056
DB01701	241.2307994
DB07738	241.3989021
DB08180	250.8631445

DB03767	264.4079981
DB04200	267.0730989
DB06969	268.9406653
DB01602	269.668359
DB03878	281.1882788
DB08499	282.1318282
DB01326	284.6668605
DB01858	285.9534691
DB07460	291.9497477
DB07177	329.6714838
DB01295	356.0202432
DB02549	361.6608731
DB08001	363.8151866
DB07735	363.8810707
DB07784	371.5522982
DB04369	373.306971
DB06959	381.9726162
DB07713	390.5724441
DB07772	390.5724441
DB03744	396.9262088
DB07853	397.2935213
DB03796	402.4330891
DB07793	425.7246941
DB01604	433.5249678
DB07657	446.9978686
DB04572	448.3758452
DB01319	459.7327498
DB06909	460.805867
DB07615	464.0013879
DB01745	465.5363892
DB06964	469.1931748
DB07572	478.1311267
DB08526	485.0925619
DB07861	498.9910201
DB02669	506.323588
DB01095	508.2611968
DB08240	514.4301996
DB04570	522.5824218
DB08354	540.0646063
DB07831	542.4471393
DB07231	543.4852472
DB08170	570.2867154
DB08463	593.5113956
DB08105	613.2776584
DB01067	624.328784
DB07750	634.5891503
DB04695	636.6225541
DB04534	760.7171982
DB07520	773.1374578
DB07786	798.6074522
DB07172	839.8626588
DB08118	847.5916728
DB01136	861.900571
DB07225	888.0008193
DB03202	890.6274184
DB07046	895.4470844
DB08640	950.2436511
DB08009	976.7401323
DB07111	983.7956596
DB03450	983.8869683
DB01819	1006.325764
DB07780	1037.084504
DB01425	1080.166986
DB08322	1147.622536
DB04049	1157.613217
DB07264	1188.577178
DB08732	1197.779373
DB07916	1246.458843
DB08026	1258.816699
DB02555	1292.74357
DB07760	1293.232883
DB04657	1414.266029
DB08762	1417.326114
DB01940	1521.981973
DB06979	1541.644238
DB03440	1635.661015
DB07209	1679.137591
DB08081	1721.804615
DB08169	1736.351234
DB08775	1741.458531
DB04732	1747.173123
DB04580	1774.818325
DB07272	1776.563286
DB00807	1813.413081
DB04179	1833.323367
DB00346	1899.679022
DB08810	1923.110062
DB07160	1953.563131
DB03110	1976.296631
DB03343	2108.265779
DB07066	2109.00372
DB08489	2134.203236
DB07113	2173.392764
DB07970	2186.707526
DB07351	2205.425021
DB06986	2224.341776
DB06993	2260.510209
DB02056	2339.274871
DB04133	2352.804051
DB03814	2377.678059
DB07771	2388.339163
DB07224	2394.685157
DB00770	2405.636207
DB01991	2406.031195
DB08275	2411.259502
DB07558	2517.084163
DB04542	2538.580034
DB07246	2597.008548
DB07934	2642.942529
DB03880	2695.329172
DB08366	2712.810619
DB08505	2714.792197
DB03359	2719.413667
DB07881	2745.110132
DB04773	2754.144915
DB08342	2774.808761
DB03249	2782.390222
DB02565	2829.382615
DB03629	2854.745764
DB07779	2927.762803
DB03434	2929.364525
DB02395	3186.941898
DB02620	3239.146823
DB08005	3325.930351
DB02095	3339.031609
DB01487	3450.826309
DB08407	3487.387072
DB01891	3819.647726
DB04751	3947.388606
DB01101	4016.007853
DB04624	4065.830416
DB08643	4065.890705
DB00938	4409.613434
DB07221	4454.15548
DB08404	4612.893553
DB01416	4671.329805
DB08116	4718.625533
DB02477	4916.975555
DB04517	4992.319191
DB04093	5174.284067
DB07571	5414.09464
DB01332	5484.139135
DB06960	5567.337805
DB02552	5681.282628
DB03735	5704.410077
DB00905	5816.222628
DB01258	5947.64739
DB00214	6252.160844
DB07792	6553.022553
DB04619	6643.717987
DB07765	6645.612324
DB00175	6808.899973
DB07089	6884.316799
DB04299	7400.736969
DB08482	7566.687237
DB08465	8175.047486
DB03793	8292.801231
DB08325	8737.264097
DB00391	8823.326637
DB07895	8900.748998
DB04553	9494.723661
DB07740	9585.688912
DB02558	9785.968516
DB07845	9892.074633
DB02733	10262.35376
DB04320	10915.41632
DB01762	11162.14763
DB06835	11539.12017
DB07475	11553.88415
DB03090	12079.63062
DB06859	12331.57269
DB00418	12614.14586
DB02979	13443.69641
DB07801	13587.64624

DB07413	13743.97802
DB02826	14881.74113
DB02263	14884.72318
DB08229	15046.45379
DB03103	16356.89987
DB04680	18065.92501
DB04592	18848.85804
DB04310	19204.42782
DB08185	19616.00924
DB02333	20192.89741
DB03029	20201.01146
DB02597	20473.21013
DB04501	20504.21444
DB08437	21132.96316
DB01578	22293.94153
DB08586	22748.89607
DB07537	23804.00609
DB02104	24807.24658
DB01612	25859.7386
DB07325	27457.83637
DB04242	27738.7725
DB02115	27913.33601
DB02855	28460.25827
DB07448	29006.04531
DB00694	29723.85962
DB08438	29918.85376
DB07926	30238.37614
DB07633	30706.23877
DB07397	32932.83197
DB03309	33508.73812
DB08375	34210.15966
DB03573	35448.97692
DB03691	35448.97692
DB03692	35448.97692
DB00447	35660.5266
DB04485	39492.30877
DB08420	42531.46759
DB07269	45700.37589
DB08464	46683.72154
DB04145	49258.21455
DB02046	49765.22333
DB08733	54129.24078
DB03077	54398.93747
DB07608	56457.14511
DB07898	57111.02633
DB01737	57897.43825
DB03178	59172.51202
DB01627	67512.01671
DB02062	75980.7839
DB08670	77823.08096
DB01920	77966.44695
DB02632	77966.44695
DB00826	81700.29935
DB07370	83595.28553
DB00983	85741.22702
DB01274	85741.22702
DB07627	86882.87684
DB01709	102920.8671
DB04590	103695.7638
DB01177	106697.6618
DB01635	109115.3673
DB08090	118556.8244
DB08669	126016.1721
DB03919	150073.1451
DB07432	158635.9683
DB03446	167248.4945
DB08403	183847.3344
DB00183	188964.8394
DB04075	194732.3898
DB03708	209496.5094
DB07041	212151.6545
DB07602	224075.4437
DB03319	227558.5416
DB04771	252140.4824
DB04386	259595.426
DB02255	259776.1199
DB03673	266858.5559
DB03732	277498.0177
DB03741	284275.197
DB03382	306379.5726
DB07379	306763.5896
DB08704	337044.4817
DB00445	343713.9126
DB00997	343713.9126
DB03550	344856.4635
DB02783	350682.4707
DB04594	369941.2246
DB03541	374848.2613
DB01883	381833.2465
DB07341	391215.312
DB03338	420733.9939
DB03730	424531.269
DB02993	459707.1295
DB04092	491100.0022
DB07946	552101.2589
DB07659	561217.5787
DB04421	566753.9966
DB04579	610477.7331
DB04622	689927.8822
DB03819	691897.6283
DB04487	694764.4511
DB04015	777201.2776
DB02301	787284.8102
DB01965	819116.1731
DB04781	834478.6733
DB03053	852005.0717
DB01683	891330.2667
DB01872	1025121.84
DB02696	1059316.23
DB00278	1090636.513
DB04139	1214614.071
DB07377	1281175.975
DB03830	1345439.326
DB02236	1523451.425
DB03459	1625062.118
DB03240	1862937.374
DB03018	1989201.54
DB02034	2100725.338
DB04441	2172473.197
DB04187	2250674.109
DB02884	2707459.306
DB02355	2730567.986
DB01660	2781364.561
DB08163	3391956.214
DB02057	3437149.437
DB02210	3498403.568
DB03133	3693962.436
DB04084	3695356.725
DB02590	3780356.965
DB01362	3898714.844
DB03564	4008827.711
DB03648	4477664.397
DB01713	4508732.679
DB02641	4896376.956
DB02907	6313241.99
DB01842	6990995.439
DB02041	8045178.219
DB01979	8446697.143
DB01812	8577475.012
DB03431	8577475.012
DB02892	9131600.757
DB02151	9725572.673
DB00646	10817056.5
DB02663	12003052.5
DB01829	21557341.64
DB02484	22044316.88
DB01716	26125077
DB01738	29144402.8
DB01864	42249103.2
DB01839	45529833.76
DB03991	46155768.4
DB03870	60360464.4
DB04036	61318536.69
DB03070	66462085.47
DB03537	80873494.34
DB01249	84462343.7
DB07105	106056323.8
DB04261	110158016.2
DB02229	115846503
DB03862	141227827.1
DB02370	152405217.8
DB04789	165074615
DB02017	166690307.7
DB02385	204139685.5
DB02065	206695425.6
DB01755	312035374.3
DB01914	588550115.4
DB04658	703651348.5
DB03618	1040030546
DB03300	1144545329
DB02697	1340252624
DB02397	1838798601
DB03344	2331013037
DB02082	2475402019

DB03321	2552621120
DB04497	2974157198
DB03943	6281636668
DB03855	6739900900
DB00914	15074052275
DB02501	17341202790
DB04425	18066310608
DB04189	27859532660
DB02787	1.00E+11
DB02349	1.28E+11
DB04276	4.35E+11
DB01922	2.91E+12
DB02237	3.82E+12

DB03754	7.86E+12
DB02717	5.94E+13
DB03971	1.53E+16
DB04626	1.86E+16
DB00225	2.16E+53
DB00644	2.16E+53
DB00803	2.16E+53
DB01135	2.16E+53
DB01160	2.16E+53
DB01226	2.16E+53
DB01287	2.16E+53
DB01361	2.16E+53
DB03393	2.16E+53

DB03995	2.16E+53
DB04156	2.16E+53
DB04219	2.16E+53
DB04498	2.16E+53
DB04529	2.16E+53
DB06699	2.16E+53
DB03906	2.01E+110
DB03934	4.36E+110
DB03189	4.93E+112
DB04231	6.48E+112
DB00364	1.96E+115
DB02937	3.60E+118
DB00014	

No. of compounds with Predicted IC50 under 1nM = 42

No. of compounds with Predicted IC50 under 10nM = 119

No. of compounds with Predicted IC50 under 100nM = 272

No. of compounds with Predicted IC50 under 1000nM = 394

No. of compounds with Predicted IC50 under 10000nM = 498

4.4 Docking study

4.4.1 Binding site definition

The binding sites were defined using QsiteFinder. The binding site with Asp 385 was identified. Its geometric co-ordinates were noted as x= -5, y=72, z=-14, X=11, Y= 86, Z=1. These define the box for binding analysis in the docking program.

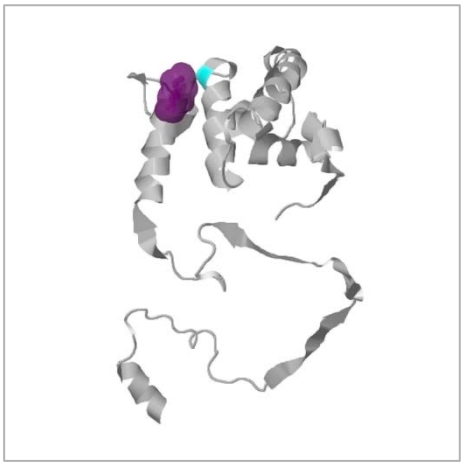

Figure 4 - Binding site with Asp 385 predicted using QsiteFinder on the CTF of presenilin (2KR6). The binding site's surface is marked in purple, while Asp 385 is marked in cyan.

4.4.2 Docking using SwissDock

Docking was performed with SwissDock. The receptor was CTF of Presenilin 2KR6 while the 498 compounds with predicted IC50 >=10000 nM, were the ligands used. The binding box was defined with the geometric co-ordinates predicted using QsiteFinder tool, x= -5, y=72, z=-14, X=11, Y= 86, Z=1. Docking study yielded 55 compounds with estimated free energy of binding lesser than -8.00 kcal/mol (Figure 5a & 5b). The range of estimated free energy of binding is -12.39 kcal/mol to -5.20 kcal/mol. The attempt to perform docking studies failed for 51 compounds either due to complexity of the ligand

structure for the docking program or due to loss of data due to the crashing of docking program.

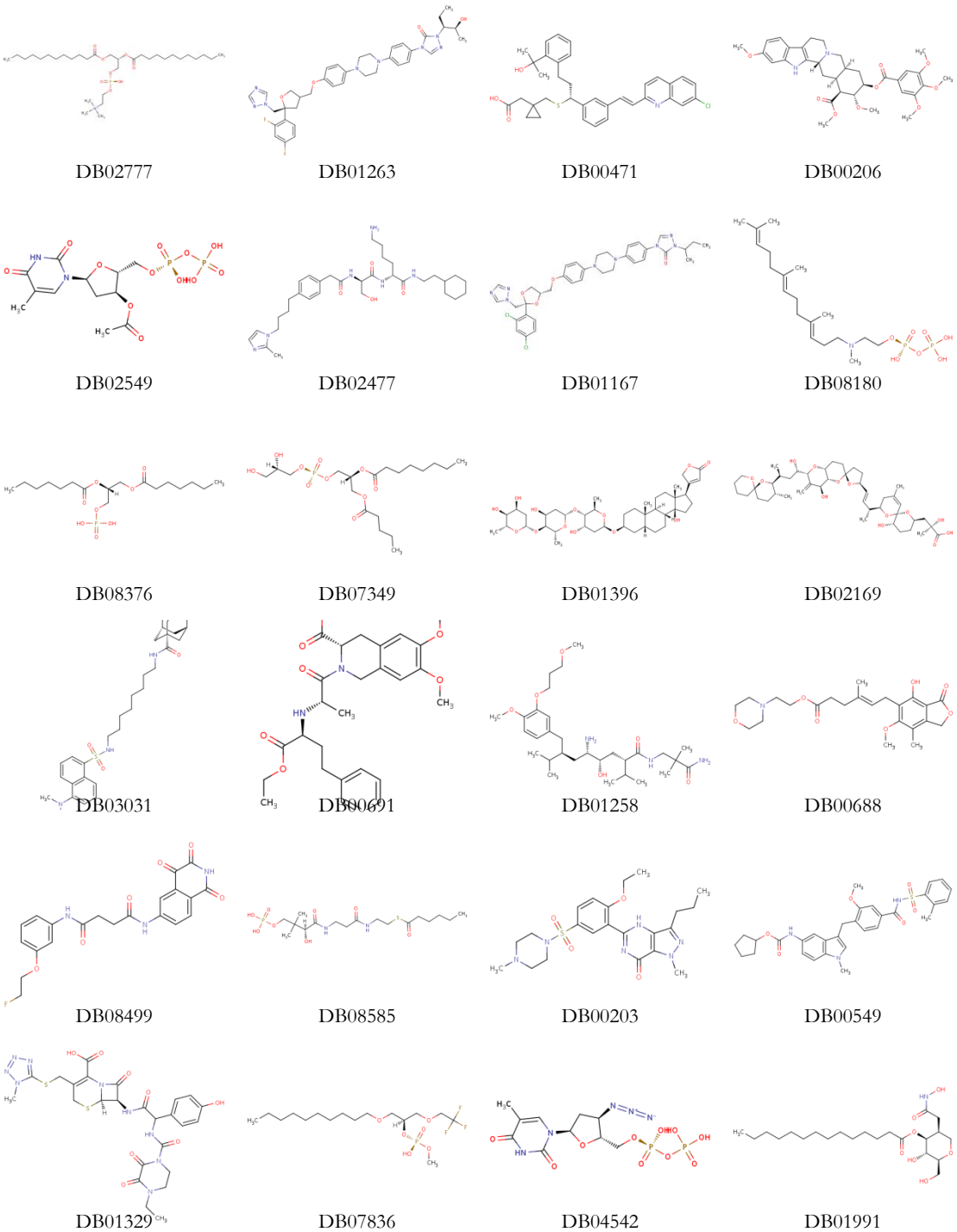

Figure 5a – Hits of docking study and estimated free energy of binding screen ΔG <= -8.00 kcal/mol for the compounds with predicted IC50 <=10000nM

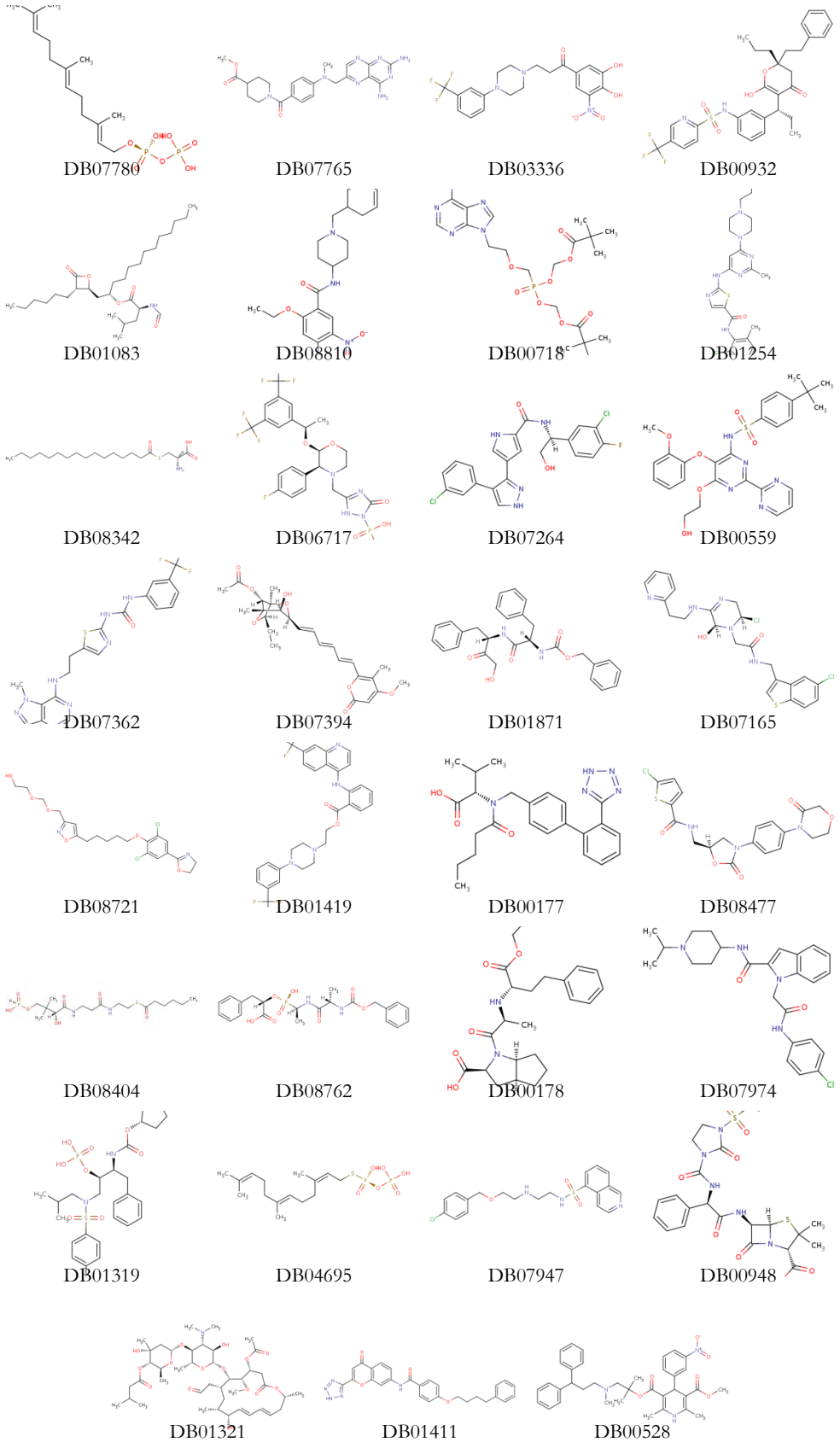

Figure 5b – Hits of docking study and estimated free energy of binding screen, ΔG <= -8.00 kcal/mol for the compounds with predicted IC50 <=10000nM

Table 8 – Result of docking study showing estimated free energy of binding for the compounds with predicted IC50 <=10000nM

DrugBank Id	ΔG (kcal/mol)
DB02777	-12.39
DB01263	-9.12
DB00471	-8.92
DB00206	-8.87
DB02549	-8.86
DB02477	-8.77
DB01167	-8.60
DB08180	-8.59
DB08376	-8.53
DB07349	-8.52
DB01396	-8.45
DB02169	-8.41
DB03031	-8.41
DB00691	-8.37
DB01258	-8.36
DB00688	-8.35
DB08499	-8.35
DB08585	-8.35
DB00203	-8.33
DB00549	-8.32
DB01329	-8.31
DB07836	-8.31
DB04542	-8.30
DB01991	-8.29
DB07780	-8.28
DB07765	-8.26
DB03336	-8.25
DB00932	-8.23
DB01083	-8.21
DB08810	-8.21
DB00718	-8.20
DB01254	-8.20
DB08342	-8.20
DB06717	-8.18
DB07264	-8.17
DB00559	-8.16
DB07362	-8.15
DB07394	-8.15
DB01871	-8.14
DB07165	-8.13
DB08721	-8.13
DB01419	-8.09
DB00177	-8.08
DB08477	-8.08
DB08404	-8.07
DB08762	-8.07
DB00178	-8.06
DB07974	-8.06
DB01319	-8.04
DB04695	-8.04
DB07947	-8.04
DB00948	-8.02
DB01321	-8.02
DB01411	-8.02
DB00528	-8.01
DB01067	-7.98
DB07675	-7.97
DB07870	-7.97
DB03556	-7.95
DB08490	-7.95
DB08429	-7.93
DB07844	-7.92
DB04270	-7.91
DB01095	-7.90
DB07255	-7.90
DB07982	-7.90
DB00251	-7.89
DB01332	-7.89
DB04645	-7.89
DB07630	-7.89
DB07995	-7.89
DB02620	-7.88
DB07068	-7.87
DB07251	-7.87
DB06960	-7.86
DB07875	-7.86
DB03984	-7.85
DB04570	-7.85
DB04642	-7.85
DB07624	-7.85
DB08710	-7.84
DB07861	-7.83
DB02804	-7.82
DB03343	-7.82
DB07221	-7.82
DB07881	-7.82
DB00619	-7.81
DB01149	-7.81
DB03359	-7.81
DB07916	-7.81
DB00773	-7.80
DB08026	-7.80
DB08487	-7.80
DB01251	-7.79
DB02545	-7.79
DB04517	-7.79
DB07369	-7.79
DB08174	-7.79
DB08749	-7.79
DB07583	-7.78
DB01326	-7.77
DB02669	-7.77
DB06829	-7.76
DB07111	-7.76
DB08755	-7.76
DB00243	-7.75
DB01264	-7.75
DB07231	-7.75
DB00346	-7.74
DB07460	-7.74
DB00661	-7.73
DB07062	-7.73
DB07333	-7.72
DB06892	-7.71
DB04609	-7.70
DB04673	-7.69
DB06844	-7.69
DB06993	-7.69
DB07750	-7.69
DB08463	-7.69
DB01342	-7.68
DB02706	-7.68
DB06846	-7.68
DB00912	-7.67
DB01098	-7.67
DB01940	-7.67
DB04632	-7.67
DB06886	-7.67
DB07272	-7.67
DB07582	-7.67
DB01180	-7.66
DB00342	-7.65
DB00391	-7.65
DB00938	-7.65
DB07334	-7.65
DB01016	-7.64
DB01162	-7.64
DB02056	-7.64
DB02253	-7.64
DB06969	-7.64
DB07691	-7.64
DB08014	-7.64
DB08173	-7.64
DB07089	-7.63
DB07832	-7.62
DB07988	-7.62
DB08078	-7.62
DB01088	-7.61
DB03249	-7.61
DB07278	-7.61
DB03367	-7.60
DB04644	-7.60
DB08275	-7.60
DB07804	-7.59
DB07839	-7.59
DB08143	-7.59
DB08325	-7.59
DB03837	-7.58
DB06909	-7.58
DB07209	-7.58
DB07786	-7.58
DB07845	-7.58
DB08177	-7.58

DB08446	-7.58
DB00451	-7.57
DB01858	-7.57
DB03785	-7.57
DB04140	-7.57
DB07351	-7.57
DB07771	-7.57
DB08116	-7.57
DB08526	-7.57
DB08722	-7.57
DB03440	-7.56
DB04133	-7.56
DB08322	-7.56
DB01029	-7.55
DB07113	-7.55
DB01745	-7.54
DB00385	-7.53
DB08341	-7.53
DB08483	-7.53
DB00950	-7.52
DB07321	-7.52
DB07769	-7.52
DB08444	-7.52
DB01349	-7.51
DB03202	-7.51
DB07169	-7.51
DB07403	-7.51
DB07571	-7.51
DB08147	-7.51
DB00444	-7.50
DB03450	-7.50
DB07967	-7.50
DB08266	-7.50
DB08349	-7.50
DB08633	-7.50
DB00796	-7.49
DB01348	-7.49
DB02848	-7.48
DB07046	-7.47
DB08590	-7.47
DB00696	-7.46
DB00770	-7.46
DB01136	-7.46
DB01324	-7.46
DB07172	-7.46
DB08152	-7.46
DB08223	-7.46
DB01347	-7.45
DB03944	-7.45
DB07853	-7.45
DB07964	-7.45
DB00607	-7.44
DB01416	-7.44
DB04739	-7.44
DB06953	-7.44
DB07558	-7.44
DB08221	-7.44
DB00481	-7.43
DB00590	-7.43
DB02558	-7.43
DB07760	-7.43
DB00562	-7.42
DB01046	-7.42
DB02071	-7.42
DB07826	-7.42
DB00622	-7.41
DB07177	-7.41
DB07740	-7.41
DB07831	-7.41
DB07848	-7.41
DB06959	-7.40
DB06974	-7.40
DB07056	-7.40
DB07936	-7.40
DB08465	-7.40
DB04751	-7.39
DB08324	-7.39
DB08727	-7.39
DB08772	-7.39
DB01751	-7.38
DB07090	-7.38
DB07219	-7.38
DB08278	-7.38
DB01295	-7.37
DB07093	-7.37
DB08481	-7.37
DB00905	-7.36
DB01138	-7.36
DB01691	-7.36
DB03767	-7.36
DB00287	-7.35
DB00901	-7.35
DB07800	-7.35
DB07895	-7.35
DB08622	-7.35
DB04234	-7.34
DB07692	-7.34
DB08505	-7.34
DB00372	-7.33
DB07223	-7.33
DB07792	-7.33
DB07970	-7.33
DB08226	-7.33
DB00541	-7.32
DB04232	-7.32
DB07246	-7.32
DB07981	-7.32
DB08706	-7.32
DB01200	-7.31
DB07359	-7.31
DB07878	-7.31
DB08719	-7.31
DB00705	-7.30
DB08388	-7.30
DB08582	-7.30
DB00802	-7.29
DB01891	-7.29
DB03890	-7.29
DB07416	-7.29
DB07784	-7.29
DB08282	-7.29
DB08344	-7.29
DB00274	-7.28
DB00836	-7.28
DB04643	-7.28
DB07541	-7.28
DB08460	-7.27
DB00762	-7.26
DB00862	-7.26
DB03744	-7.26
DB08195	-7.26
DB03796	-7.25
DB04416	-7.25
DB06970	-7.25
DB07250	-7.25
DB08154	-7.25
DB01604	-7.23
DB08664	-7.22
DB08732	-7.22
DB07224	-7.21
DB07429	-7.21
DB07640	-7.21
DB08720	-7.21
DB00457	-7.20
DB02550	-7.20
DB07225	-7.20
DB00502	-7.19
DB00881	-7.19
DB00910	-7.19
DB04049	-7.19
DB04200	-7.19
DB08169	-7.19
DB08310	-7.19
DB08471	-7.18
DB01089	-7.17
DB02932	-7.17
DB04505	-7.17
DB07847	-7.17
DB08458	-7.17
DB01118	-7.16
DB01128	-7.16
DB01462	-7.16
DB07471	-7.16
DB07249	-7.15
DB07615	-7.15
DB08387	-7.15
DB04012	-7.14
DB07572	-7.14
DB08407	-7.14
DB08457	-7.14
DB08459	-7.14
DB02565	-7.13
DB06961	-7.13
DB04624	-7.12
DB07188	-7.12
DB07827	-7.12
DB08366	-7.12

DB08489	-7.12
DB08556	-7.12
DB08811	-7.12
DB06979	-7.10
DB00320	-7.09
DB01388	-7.09
DB03383	-7.09
DB07779	-7.09
DB00767	-7.08
DB01039	-7.08
DB01340	-7.08
DB07962	-7.08
DB01523	-7.07
DB07738	-7.07
DB07770	-7.07
DB04574	-7.06
DB08640	-7.06
DB00317	-7.05
DB03814	-7.05
DB07713	-7.05
DB00929	-7.04
DB03878	-7.04
DB08435	-7.04
DB00682	-7.02
DB06964	-7.02
DB08482	-7.02
DB07309	-7.01
DB08118	-7.01
DB01148	-7.00
DB08354	-6.99
DB00587	-6.98
DB07735	-6.98
DB07793	-6.98
DB00680	-6.97
DB02894	-6.97
DB07943	-6.97
DB08502	-6.97
DB01459	-6.96
DB04746	-6.96
DB06986	-6.96
DB08350	-6.92
DB00527	-6.91
DB07778	-6.91
DB08614	-6.90
DB00973	-6.89
DB07066	-6.89
DB07276	-6.89
DB01570	-6.88
DB08775	-6.87
DB08081	-6.86
DB01712	-6.85
DB07846	-6.84
DB00439	-6.83
DB00882	-6.83
DB07724	-6.83
DB01101	-6.81
DB01256	-6.79
DB01464	-6.79
DB03434	-6.79
DB03735	-6.79
DB01487	-6.78
DB01425	-6.77
DB01602	-6.73
DB00807	-6.71
DB08170	-6.71
DB01070	-6.69
DB02946	-6.66
DB08302	-6.65
DB08105	-6.64
DB01505	-6.63
DB07446	-6.63
DB07772	-6.62
DB08379	-6.60
DB08690	-6.57
DB01134	-6.55
DB04773	-6.52
DB07160	-6.52
DB01892	-6.51
DB02791	-6.51
DB04317	-6.50
DB04674	-6.48
DB00596	-6.44
DB01130	-6.42
DB01819	-6.39
DB04572	-6.35
DB04693	-6.35
DB03520	-6.28
DB04657	-6.26
DB00261	-6.23
DB01410	-6.14
DB00976	-6.08
DB04534	-6.05
DB04093	-6.04
DB03554	-6.01
DB02095	-5.98
DB04580	-5.96
DB03793	-5.80
DB04299	-5.79
DB03110	-5.73
DB02721	-5.66
DB04553	-5.64
DB04619	-5.64
DB02922	-5.49
DB01701	-5.44
DB04179	-5.43
DB00732	-5.20

No. of compounds with estimated free energy of binding (<= -8.00 kcal/mol) = 55

5 Discussion

Screening of 6160 compounds from DrugBank database using various filters namely pharmacophore search, predicted biological activity (IC50) and docking study was carried out. At the end of each screening a significant amount of compounds of non-interest were eliminated from the overall screening process. There were 721 compounds after application of pharmacophore screen. After the application of biological activity screen 498 compounds were taken for further screening. These compounds were screened based on the estimated free energy of binding (ΔG) after being subjected to docking study.Finally, 54 compounds were left. These compounds belonged to various classes namely Phenylpropenes, Phenethylamines, Anisoles, Alkaloids and Alkaloid Derivatives, Steroids and Steroid derivatives, Phenylpropylamines, Salicylates and Derivatives, Benzenesulfonamides, Pyrans, Lactones, Benzamide, Purine and purine derivatives, Piperazines, Halobenzenes, Biphenyltetrazoles and Derivatives, Aminoglycosides, Chromones and Diphenylmethanes. Of these classes of compounds a few have already been associated with inhibition of gamma-secretase namely sulfonamides (61, 98, 99, 109). The occurrence of this class of compounds in the study reaffirms the potential of these compounds. Also a few of the compounds classes have not yet been reported as potential inhibitors of gamma secretase namely piperazine, phenylpropylamine, and anisoles. Also a few of the hits are anti-hypertensive drugs. This class of drugs has already been studied for therapy of Parkinson's disease(116). Despite the differing mechanisms of pathogenesis of Alzhiemer's disease and Parkinson's disease, it is inferred that further study of anit-hypertensive drugs for development of therapy of Alzheimer's may produce potential drugs. Further study of these classes of compounds may produce a novel class of inhibitors of gamma secretase. Thus it may aid in the therapy of Alzheimer's disease.

6 References

1. Dobson CM. Protein folding and misfolding. Nature. 2003;426(6968):884-90.

2. Dobson CM, editor. Principles of protein folding, misfolding and aggregation. Seminars in cell & developmental biology; 2004: Elsevier.

3. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins: Structure, Function, and Bioinformatics. 1995;21(3):167-95.

4. Dobson CM. The structural basis of protein folding and its links with human disease. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences. 2001;356(1406):133-45.

5. Bukau B, Horwich AL, Cycles H. The Hsp70 and Hsp60 review chaperone machines. Cell. 1998;92:351-66.

6. Schiene C, Fischer G. Enzymes that catalyse the restructuring of proteins. Current opinion in structural biology. 2000;10(1):40-5.

7. Matouschek A. Protein unfolding—an important process< i> in vivo</i>? Current opinion in structural biology. 2003;13(1):98-109.

8. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R, et al. Structure of the cross-beta spine of amyloid-like fibrils. Nature. 2005;435(7043):773-8. Epub 2005/06/10.

9. Linding R, Schymkowitz J, Rousseau F, Diella F, Serrano L. A comparative study of the relationship between protein structure and b-aggregation in globular and intrinsically disordered proteins. J Mol Biol. 2004;342:345-53.

10. Astbury WT, Dickinson S, Bailey K. The X-ray interpretation of denaturation and the structure of the seed globulins. Biochemical Journal. 1935;29(10):2351.

11. Jenkins J, Pickersgill R. The architecture of parallel< i> β</i>-helices and related folds. Progress in biophysics and molecular biology. 2001;77(2):111-75.

12. Roher AE, Baudry J, Chaney MO, Kuo Y-M, Stine WB, Emmerling MR. Oligomerization and fibril assembly of the amyloid-β protein. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease. 2000;1502(1):31-43.

13. Nilsberth C, Westlind-Danielsson A, Eckman CB, Condron MM, Axelman K, Forsell C, et al. The'Arctic'APP mutation (E693G) causes Alzheimer's disease by enhanced Aβ protofibril formation. Nature neuroscience. 2001;4(9):887-93.

14. Silveira JR, Raymond GJ, Hughson AG, Race RE, Sim VL, Hayes SF, et al. The most infectious prion protein particles. Nature. 2005;437(7056):257-61.

15. Reixach N, Deechongkit S, Jiang X, Kelly JW, Buxbaum JN. Tissue damage in the amyloidoses: Transthyretin monomers and nonnative oligomers are the major cytotoxic species in tissue culture. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(9):2817-22.

16. Yoshiike Y, Akagi T, Takashima A. Surface structure of amyloid-β fibrils contributes to cytotoxicity. Biochemistry. 2007;46(34):9805-12.

17. Pepys M. Pathogenesis, diagnosis and treatment of systemic amyloidosis. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences. 2001;356(1406):203-11.

18. Prince M, Jackson J. World Alzheimer Report 2009: Alzheimer's Disease International; 2009.

19. Hardy J, Allsop D. Amyloid deposition as the central event in the aetiology of Alzheimer's disease. Trends in pharmacological sciences. 1991;12:383-8.

20. Selkoe DJ. The molecular pathology of Alzheimer's disease. Neuron. 1991;6:487-98.

21. Taguchi K, Yamagata HD, Zhong W, Kamino K, Akatsu H, Hata R, et al. Identification of hippocampus-related candidate genes for Alzheimer's disease. Annals of neurology. 2005;57(4):585-8.

22. Blennow K, de Leon MJ, Zetterberg H. Alzheimer's disease. Lancet. 2006;368(9533):387-403. Epub 2006/08/01.

23. Lambon Ralph MA, Patterson K, Graham N, Dawson K, Hodges JR. Homogeneity and heterogeneity in mild cognitive impairment and Alzheimer's disease: a cross-sectional and longitudinal study of 55 cases. Brain: a journal of neurology. 2003;126(Pt 11):2350.

24. Artero S, Tierney MC, Touchon J, Ritchie K. Prediction of transition from cognitive impairment to senile dementia: a prospective, longitudinal study. Acta Psychiatrica Scandinavica. 2003;107(5):390-3.

25. Huang X, Atwood CS, Hartshorn MA, Multhaup G, Goldstein LE, Scarpa RC, et al. The A beta peptide of Alzheimer's disease directly produces hydrogen peroxide through metal ion reduction. Biochemistry. 1999;38(24):7609-16. Epub 1999/07/01.

26. Huang X, Cuajungco MP, Atwood CS, Hartshorn MA, Tyndall JD, Hanson GR, et al. Cu(II) potentiation of alzheimer abeta neurotoxicity. Correlation with cell-free hydrogen peroxide production and metal reduction. The Journal of biological chemistry. 1999;274(52):37111-6. Epub 1999/12/22.

27. Opazo C, Huang X, Cherny RA, Moir RD, Roher AE, White AR, et al. Metalloenzyme-like activity of Alzheimer's disease beta-amyloid. Cu-dependent catalytic conversion of dopamine, cholesterol, and biological reducing agents to neurotoxic H(2)O(2). The Journal of biological chemistry. 2002;277(43):40302-8. Epub 2002/08/23.

28. Lustbader JW, Cirilli M, Lin C, Xu HW, Takuma K, Wang N, et al. ABAD directly links Abeta to mitochondrial toxicity in Alzheimer's disease. Science (New York, NY). 2004;304(5669):448-52. Epub 2004/04/17.

29. Tuppo EE, Arias HR. The role of inflammation in Alzheimer's disease. The international journal of biochemistry & cell biology. 2005;37(2):289-305.

30. Weldon DT, Rogers SD, Ghilardi JR, Finke MP, Cleary JP, O'Hare E, et al. Fibrillar β-amyloid induces microglial phagocytosis, expression of inducible nitric oxide synthase, and loss of a select population of neurons in the rat CNS in vivo. The Journal of neuroscience. 1998;18(6):2161-73.

31. Guillemin GJ, Williams KR, Smith DG, Smythe GA, Croitoru-Lamoury J, Brew BJ. Quinolinic acid in the pathogenesis of Alzheimer's disease. Developments in Tryptophan and Serotonin Metabolism: Springer; 2003. p. 167-76.

32. Hardy J. An 'anatomical cascade hypothesis' for Alzheimer's disease. Trends in neurosciences. 1992;15(6):200-1.

33. De Strooper B, Umans L, Van Leuven F, Van Den Berghe H. Study of the synthesis and secretion of normal and artificial mutants of murine amyloid precursor protein (APP): cleavage of APP occurs in a late compartment of the default secretion pathway. The Journal of cell biology. 1993;121(2):295-304.

34. Kojro E, Fahrenholz F. The non-amyloidogenic pathway: structure and function of α-secretases. Alzheimer's Disease: Springer; 2005. p. 105-27.

35. Cuajungco MP, Fagét KY. Zinc takes the center stage: its paradoxical role in Alzheimer's disease. Brain Research Reviews. 2003;41(1):44-56.

36. Kirkitadze MD, Kowalska A. Molecular mechanisms initiating amyloid beta-fibril formation in Alzheimer's disease. ACTA BIOCHIMICA POLONICA-ENGLISH EDITION-. 2005;52(2):417.

37. Dahlgren KN, Manelli AM, Stine WB, Jr., Baker LK, Krafft GA, LaDu MJ. Oligomeric and fibrillar species of amyloid-beta peptides differentially affect neuronal viability. The Journal of biological chemistry. 2002;277(35):32046-53. Epub 2002/06/12.

38. Levy-Lahad E, Bird TD. Genetic factors in Alzheimer's disease: a review of recent advances. Annals of neurology. 1996;40(6):829-40.

39. Capell A, Steiner H, Romig H, Keck S, Baader M, Grim MG, et al. Presenilin-1 differentially facilitates endoproteolysis of the &bgr;-amyloid precursor protein and Notch. Nature cell biology. 2000;2(4):205-11.

40. Kowalska A. Genetic basis of neurodegeneration in familial Alzheimer's disease. Pol J Pharmacol. 2004;56(2):171-8.

41. Cruts M, Theuns J, Van Broeckhoven C. Locus-specific mutation databases for neurodegenerative brain diseases. Human mutation. 2012;33(9):1340-4. Epub 2012/05/15.

42. Wolfe MS, Xia W, Ostaszewski BL, Diehl TS, Kimberly WT, Selkoe DJ. Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and γ-secretase activity. Nature. 1999;398(6727):513-7.

43. Thinakaran G, Borchelt DR, Lee MK, Slunt HH, Spitzer L, Kim G, et al. Endoproteolysis of presenilin 1 and accumulation of processed derivatives in vivo. Neuron. 1996;17(1):181-90.

44. Edbauer D, Winkler E, Haass C, Steiner H. Presenilin and nicastrin regulate each other and determine amyloid β-peptide production via complex formation. Proceedings of the National Academy of Sciences. 2002;99(13):8666-71.

45. Goutte C, Tsunozaki M, Hale VA, Priess JR. APH-1 is a multipass membrane protein essential for the Notch signaling pathway in Caenorhabditis elegans embryos. Proceedings of the National Academy of Sciences. 2002;99(2):775-9.

46. Francis R, McGrath G, Zhang J, Ruddy DA, Sym M, Apfeld J, et al. aph-1 and pen-2 are required for Notch pathway signaling, gamma-secretase cleavage of betaAPP, and presenilin protein accumulation. Developmental cell. 2002;3(1):85.

47. Kimberly WT, LaVoie MJ, Ostaszewski BL, Ye W, Wolfe MS, Selkoe DJ. γ-Secretase is a membrane protein complex comprised of presenilin, nicastrin, aph-1, and pen-2. Proceedings of the National Academy of Sciences. 2003;100(11):6382-7.

48. Edbauer D, Winkler E, Regula JT, Pesold B, Steiner H, Haass C. Reconstitution of γ-secretase activity. Nature cell biology. 2003;5(5):486-8.

49. Hayashi I, Urano Y, Fukuda R, Isoo N, Kodama T, Hamakubo T, et al. Selective reconstitution and recovery of functional γ-secretase complex on budded baculovirus particles. Journal of Biological Chemistry. 2004;279(36):38040-6.

50. Takasugi N, Tomita T, Hayashi I, Tsuruoka M, Niimura M, Takahashi Y, et al. The role of presenilin cofactors in the γ-secretase complex. Nature. 2003;422(6930):438-41.

51. Searfoss GH, Jordan WH, Calligaro DO, Galbreath EJ, Schirtzinger LM, Berridge BR, et al. Adipsin, a biomarker of gastrointestinal toxicity mediated by a functional γ-secretase inhibitor. Journal of Biological Chemistry. 2003;278(46):46107-16.

52. Lleó A, Berezovska O, Herl L, Raju S, Deng A, Bacskai BJ, et al. Nonsteroidal anti-inflammatory drugs lower Aβ42 and change presenilin 1 conformation. Nature medicine. 2004;10(10):1065-6.

53. Shearman MS, Beher D, Clarke EE, Lewis HD, Harrison T, Hunt P, et al. L-685,458, an aspartyl protease transition state mimic, is a potent inhibitor of amyloid β-protein precursor γ-secretase activity. Biochemistry. 2000;39(30):8698-704.

54. Boggs LN, Fuson KS, Gitter BD, Czilli DL, Hyslop PA, Bender MH, et al. P1-419 In vivo characterization of LY450139, a novel, stereoselective, functional gamma-secretase inhibitor. Neurobiology of aging. 2004;25:S218-S.

55. Lanz TA, Himes CS, Pallante G, Adams LA, Yamazaki S, Amore BM, et al. The γ-secretase inhibitor, DAPT, reduces Aβ levels in vivo in plasma and CSF in young (plaque-free) and aged (plaque-bearing) Tg2576 mice. Journal of Pharmacology and Experimental Therapeutics. 2003.

56. Netzer WJ, Dou F, Cai D, Veach D, Jean S, Li Y, et al. Gleevec inhibits β-amyloid production but not Notch cleavage. Proceedings of the National Academy of Sciences. 2003;100(21):12444-9.

57. Takahashi Y, Hayashi I, Tominari Y, Rikimaru K, Morohashi Y, Kan T, et al. Sulindac sulfide is a noncompetitive γ-secretase inhibitor that preferentially reduces Aβ42 generation. Journal of Biological Chemistry. 2003;278(20):18664-70.

58. FORSBLOM R, PAULSEN K, ROTTICCI D, SANTANGELO E, WALDMAN M. MODULATORS OF AMYLOID BETA. WO Patent 2,010,053,438; 2010.

59. Baumann K, Goetschi E, Green L, Jolidon S, Knust H, Limberg A, et al. GAMMA SECRETASE MODULATORS. Google Patents; 2011.

60. Wallace OB, Smith DW, Deshpande MS, Polson C, Felsenstein KM. Inhibitors of Abeta production: solid-phase synthesis and SAR of alpha-hydroxycarbonyl derivatives. Bioorganic & medicinal chemistry letters. 2003;13(6):1203-6. Epub 2003/03/20.

61. Teall M, Oakley P, Harrison T, Shaw D, Kay E, Elliott J, et al. Aryl sulfones: a new class of gamma-secretase inhibitors. Bioorganic & medicinal chemistry letters. 2005;15(10):2685-8. Epub 2005/05/03.

62. Lewis SJ, Smith AL, Neduvelil JG, Stevenson GI, Lindon MJ, Jones AB, et al. A novel series of potent gamma-secretase inhibitors based on a benzobicyclo[4.2.1]nonane core. Bioorganic & medicinal chemistry letters. 2005;15(2):373-8. Epub 2004/12/18.

63. Bihel F, Das C, Bowman MJ, Wolfe MS. Discovery of a Subnanomolar helical D-tridecapeptide inhibitor of gamma-secretase. J Med Chem. 2004;47(16):3931-3. Epub 2004/07/23.

64. Findeis MA, Schroeder F, McKee TD, Yager D, Fraering PC, Creaser SP, et al. Discovery of a novel pharmacological and structural class of gamma secretase modulators derived from the extract of actaea racemosa. ACS chemical neuroscience. 2012;3(11):941-51.

65. Hubbs JL, Fuller NO, Austin WF, Shen R, Creaser SP, McKee TD, et al. Optimization of a Natural Product-Based Class of γ-Secretase Modulators. Journal of Medicinal Chemistry. 2012;55(21):9270-82.

66. Huang X, Zhou W, Liu X, Li H, Sun G, Mandal M, et al. Synthesis and SAR Studies of Fused Oxadiazines as γ-Secretase Modulators for Treatment of Alzheimer's Disease. ACS Medicinal Chemistry Letters. 2012;3(11):931-5.

67. Crump CJ, Fish BA, Castro SV, Chau D-M, Gertsik N, Ahn K, et al. Piperidine acetic acid based γ-secretase modulators directly bind to presenilin-1. ACS chemical neuroscience. 2011;2(12):705-10.

68. Miller MJH, Porter WJ, Reel JK, Rubio-Esteban A. Azepine derivatives as gamma-secretase inhibitors. Google Patents; 2012.

69. Takahashi Y, Fuwa H, Kaneko A, Sasaki M, Yokoshima S, Koizumi H, et al. Novel gamma-secretase inhibitors discovered by library screening of in-house synthetic natural product intermediates. Bioorganic & medicinal chemistry letters. 2006;16(14):3813-6. Epub 2006/05/10.

70. Walters WP, Stahl MT, Murcko MA. Virtual screening--an overview. Drug discovery today. 1998;3(4):160-78.

71. Fara DC, Oprea TI, Prossnitz ER, Bologa CG, Edwards BS, Sklar LA. Integration of virtual and physical screening. Drug Discovery Today: Technologies. 2007;3(4):377-85.

72. Rollinger J, Langer T, Stuppner H. Strategies for efficient lead structure discovery from natural products. Current medicinal chemistry. 2006;13(13):1491-507.

73. Lyne PD. Structure-based virtual screening: an overview. Drug discovery today. 2002;7(20):1047-55.

74. Oprea TI, Matter H. Integrating virtual screening in lead discovery. Current opinion in chemical biology. 2004;8(4):349-58.

75. Nettles JH, Jenkins JL, Bender A, Deng Z, Davies JW, Glick M. Bridging chemical and biological space:"target fishing" using 2D and 3D molecular descriptors. Journal of Medicinal Chemistry. 2006;49(23):6802-10.

76. Bissantz C, Schalon C, Guba W, Stahl M. Focused library design in GPCR projects on the example of 5-HT2c agonists: Comparison of structure-based virtual screening with ligand-based search methods. Proteins: Structure, Function, and Bioinformatics. 2005;61(4):938-52.

77. Amaning K, Lowinski M, Vallee F, Steier V, Marcireau C, Ugolini A, et al. The use of virtual screening and differential scanning fluorimetry for the rapid identification of fragments active against MEK1. Bioorganic & medicinal chemistry letters. 2013.

78. Gallastegui E, Marshall B, Vidal D, Sanchez-Duffhues G, Collado JA, Alvarez-Fernández C, et al. Combination of biological screening in a cellular model of viral latency and virtual screening identifies novel compounds that reactivate HIV-1. Journal of virology. 2012;86(7):3795-808.

79. Ehrlich P. Über den jetzigen Stand der Chemotherapie. Berichte der deutschen chemischen Gesellschaft. 1909;42(1):17-47.

80. Wermuth C-G, Ganellin CR, Lindberg P, Mitscher LA. . Glossary of Terms Used in Medicinal Chemistry (IUPAC Recommendations 1997. Annual reports in medicinal chemistry. 1998;33:385-95.

81. Jones G, Willet P. GASP: genetic algorithm superimposition program. Pharmacophore Perception, Development, and Use in Drug Design (Güner, OF, ed). 2000:85-106.

82. Martin YC. DISCO: what we did right and what we missed. Pharmacophore Perception, Development, and Use in Drug Design La Jolla, CA, International University Line. 2000:51-66.

83. Barnum D, Greene J, Smellie A, Sprague P. Identification of common functional configurations among molecules. Journal of chemical information and computer sciences. 1996;36(3):563-71.

84. Li H, Sutter J, Hoffmann R. HypoGen: An automated system for generating 3D predictive pharmacophore models. Pharmacophore Perception, Development, and Use in Drug Design. 2000;2:171.

85. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. Journal of computer-aided molecular design. 2006;20(10-11):647-71.

86. Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. Journal of chemical information and modeling. 2005;45(1):160-9.

87. Chen J, Lai L. Pocket v. 2: further developments on receptor-based pharmacophore modeling. Journal of chemical information and modeling. 2006;46(6):2684-91.

88. Ortuso F, Langer T, Alcaro S. GBPM: GRID-based pharmacophore model: concept and application studies to protein–protein recognition. Bioinformatics. 2006;22(12):1449-55.

89. Blaney JM, Hansch C, Silipo C, Vittoria A. Structure-activity relationships of dihydrofolated reductase inhibitors. Chemical Reviews. 1984;84(4):333-407.

90. Ewing TJ, Makino S, Skillman AG, Kuntz ID. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. Journal of computer-aided molecular design. 2001;15(5):411-28.

91. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. Journal of molecular biology. 1996;261(3):470-89.

92. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. Journal of molecular biology. 1997;267(3):727-48.

93. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of Medicinal Chemistry. 2004;47(7):1739-49.

94. Totrov M, Abagyan R. Flexible protein-ligand docking by global energy optimization in internal coordinates. Proteins Structure Function and Genetics. 1997;29(s 1):215-20.

95. McGann MR, Almond HR, Nicholls A, Grant JA, Brown FK. Gaussian docking functions. Biopolymers. 2003;68(1):76-90.

96. Goodsell DS, Lauble H, Stout CD, Olson AJ. Automated docking in crystallography: analysis of the substrates of aconitase. Proteins: Structure, Function, and Bioinformatics. 1993;17(1):1-10.

97. Mizutani MY, Itai A. Efficient method for high-throughput virtual screening based on flexible docking: discovery of novel acetylcholinesterase inhibitors. Journal of Medicinal Chemistry. 2004;47(20):4818-28.

98. Jelley RA, Elliott J, Gibson KR, Harrison T, Beher D, Clarke EE, et al. 3-Substituted gem-cyclohexane sulfone based gamma-secretase inhibitors for Alzheimer's disease: conformational analysis and biological activity. Bioorganic & medicinal chemistry letters. 2006;16(14):3839-42. Epub 2006/05/10.

99. Sparey T, Beher D, Best J, Biba M, Castro JL, Clarke E, et al. Cyclic sulfamide gamma-secretase inhibitors. Bioorganic & medicinal chemistry letters. 2005;15(19):4212-6. Epub 2005/08/02.

100. Crump CJ, Fish BA, Castro SV, Chau DM, Gertsik N, Ahn K, et al. Piperidine acetic acid based gamma-secretase modulators directly bind to Presenilin-1. ACS chemical neuroscience. 2011;2(12):705-10. Epub 2012/01/10.

101. Wolfe MS. gamma-Secretase inhibitors and modulators for Alzheimer's disease. Journal of neurochemistry. 2012;120 Suppl 1:89-98. Epub 2011/11/30.

102. Tang BL, Liou YC. Novel modulators of amyloid-beta precursor protein processing. Journal of neurochemistry. 2007;100(2):314-23. Epub 2007/01/24.

103. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic acids research. 2012;40(D1):D1100-D7.

104. Oehlrich D, Berthelot DJ, Gijsen HJ. gamma-Secretase modulators as potential disease modifying anti-Alzheimer's drugs. J Med Chem. 2011;54(3):669-98. Epub 2010/12/15.

105. Lubbers T, Flohr A, Jolidon S, David-Pierson P, Jacobsen H, Ozmen L, et al. Aminothiazoles as gamma-secretase modulators. Bioorganic & medicinal chemistry letters. 2011;21(21):6554-8. Epub 2011/09/20.

106. Gamerdinger M, Clement AB, Behl C. Effects of sulindac sulfide on the membrane architecture and the activity of gamma-secretase. Neuropharmacology. 2008;54(6):998-1005. Epub 2008/03/25.

107. Imbimbo BP, Giardino L, Sivilia S, Giuliani A, Gusciglio M, Pietrini V, et al. CHF5074, a novel gamma-secretase modulator, restores hippocampal neurogenesis potential and reverses contextual memory deficit in a transgenic mouse model of Alzheimer's disease. Journal of Alzheimer's disease : JAD. 2010;20(1):159-73. Epub 2010/02/19.

108. Ohki Y, Higo T, Uemura K, Shimada N, Osawa S, Berezovska O, et al. Phenylpiperidine-type gamma-secretase modulators target the transmembrane domain 1 of presenilin 1. The EMBO journal. 2011;30(23):4815-24. Epub 2011/10/18.

109. Borgegard T, Jureus A, Olsson F, Rosqvist S, Sabirsh A, Rotticci D, et al. First and second generation gamma-secretase modulators (GSMs) modulate amyloid-beta (Abeta)

peptide production through different mechanisms. The Journal of biological chemistry. 2012;287(15):11810-9. Epub 2012/02/16.

110. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. Journal of cheminformatics. 2011;3(1):1-14.

111. Discovery Studio 2.5. 2.5 ed: Accelrys, Inc., San Diego, CA; 2005 - 2009.

112. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic acids research. 2006;34(suppl 1):D668-D72.

113. Laurie AT, Jackson RM. Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites. Bioinformatics. 2005;21(9):1908-16.

114. Grosdidier A, Zoete V, Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic acids research. 2011;39(suppl 2):W270-W7.

115. Grosdidier A, Zoete V, Michielin O. EADock: Docking of small molecules into protein active sites with a multiobjective evolutionary optimization. Proteins: Structure, Function, and Bioinformatics. 2007;67(4):1010-25.

116. Rees K, Stowe R, Patel S, Ives N, Breen K, Ben-Shlomo Y, et al. Anti-hypertensive drugs as disease-modifying agents for Parkinson's disease: evidence from observational studies and clinical trials. The Cochrane Library. 2010.